НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ
Кафедра інформаційної безпеки

«До захисту допущено»
В.о. завідувача кафедри

_____ М.В.Грайворонський
(підпис)
“____”_____2019 р.

Дипломна робота

на здобуття ступеня бакалавра

з напряму підготовки 6.040301 «Прикладна математика»

на тему: Системний аналіз категорії “загроза” в кібербезпеці

Виконав (-ла): студент (-ка) __4__ курсу, групи ___ФІ-51__
(шифр групи)

Яскал Назар Олександрович _____ _____
(прізвище, ім’я, по батькові) (підпис)

Керівник _____ _____
(посада, науковий ступінь, вчене звання, прізвище та ініціали) (підпис)

Консультант _____ _____ _____
(назва розділу) (посада, вчене звання, науковий ступінь, прізвище, ініціали) (підпис)

Рецензент _____ _____
(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали) (підпис)

Засвідчую, що у цій дипломній роботі немає запозичень з праць інших авторів без відповідних посилань.
Студент _____
(підпис)

Київ - 2019 року

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ
Кафедра інформаційної безпеки

Рівень вищої освіти – перший (бакалаврський)

Напрям підготовки 6.040301 «Прикладна математика»

ЗАТВЕРДЖУЮ
В.о. завідувача кафедри
_____ М.В.Грайворонський
(підпис)
«___»_____2019 р.

ЗАВДАННЯ
на дипломну роботу студенту

_____
(прізвище, ім'я, по батькові)

1. Тема роботи _____
_____
_____ ,
науковий керівник роботи _____
_____,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «____ » 2019 р. № _____

2. Термін подання студентом роботи 10 червня 2019 р.

3. Вихідні дані до роботи _____
_____
_____

4. Зміст роботи_____
_____
_____
_____
_____

5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо) _____
_____
_____
_____

6. Дата видачі завдання _____

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів дипломної роботи	Примітка

Студент _____ _____
(підпис) (ініціали, прізвище)

Керівник роботи _____ _____
(підпис) (ініціали, прізвище)

РЕФЕРАТ

Кваліфікаційна робота містить: 64 сторінки, 32 рисунків, 13 таблиць та 42 джерела.

У цій роботі був зроблений детальний аналіз термінологічної проблеми у сфері “Кібербезпеки” та на основі нього запропоновані свої визначення ключових елементів тезаурусу. Проаналізовано динаміку використання визначених термінів у наукових роботах та проведено порівняльний аналіз отриманих результатів з аналогічними результатами для відповідних термінів у сфері інформаційної безпеки.

Темою роботи є системний аналіз категорії “загроза” в кібербезпеці.

Метою даної роботи є вирішення термінологічної проблеми у сфері кібербезпеки та чітке розділення інформаційної та кібернетичної сфер.

Об’єктом дослідження є вживання зазначеної термінології у наукових та нормативно-правових документах.

Предметом дослідження є контекст в якому використовуються визначені елементи та динаміка обсягу їх використання.

Методами дослідження є інтелектуальний аналіз тексту(text mining), розвідковий аналіз, порівняльний аналіз, RS-аналіз.

Результати цієї роботи були частково представлені на XVII Науково-практичній конференції студентів, аспірантів та молодих вчених "Теоретичні і прикладні проблеми фізики, математики та інформатики"(26-27 квітня 2019р., м. Київ).

КІБЕРПРОСТІР, КІБЕРБЕЗПЕКА, КІБЕРЗАГРОЗА, КІБЕРНЕТИЧНИЙ РИЗИК, ТЕРМІНОЛОГІЯ, TEXT MINING, RS-АНАЛІЗ

РЕФЕРАТ

Квалификационная работа содержит: 64 страницы, 32 рисунков, 13 таблиц и 42 источников.

В этой работе был сделан подробный анализ терминологической проблемы в сфере "Кибербезопасности" и на основе него предложены свои определения ключевых элементов тезауруса. Проанализирована динамика использования определенных терминов в научных работах и проведен сравнительный анализ полученных результатов с аналогичными результатами для соответствующих терминов в сфере информационной безопасности.

Темой работы является системный анализ категории "угроза" в кибербезопасности.

Целью данной работы является решение терминологической проблемы в сфере кибербезопасности и четкое разделение информационной и кибернетической сфер.

Объектом исследования является употребление указанной терминологии в научных и нормативно-правовых документах.

Предметом исследования является контекст в котором используются определенные элементы и динамика объема их использования.

Методами исследования являются интеллектуальный анализ текста (text mining), разведывательно анализ, сравнительный анализ, RS-анализ.

Результаты этой работы были частично представлены на XVII Научно-практической конференции студентов, аспирантов и молодых ученых "Теоретические і прикладные проблемы физики, математики и информатики" (26-27 апреля 2019г., г. Киев).

КИБЕРПРОСТРАНСТВО, КИБЕРБЕЗОПАСНОСТЬ, КИБЕРУГРОЗА, КИБЕРНЕТИЧЕСКИЙ РИСК, ТЕРМИНОЛОГИЯ, TEXT MINING, RS-АНАЛИЗ

ABSTRACT

The thesis contains: 64 pages, 32 figures, 13 tables and 42 sources.

In this work, a detailed analysis of the terminology problem in the field of "Cyber Security" was made and based on its results proposed definitions of key elements of the thesaurus. The dynamics of the use of definite terms in scientific works is analyzed and a comparative analysis of the obtained results with similar results for the corresponding terms in the field of information security has been carried out.

The theme of the work is a Systematic Analysis of the "Threat" Category in Cybersecurity.

The goal of this work is to solve the terminological problem in the field of cybersecurity and to clearly divide the information and cybernetic spheres.

The object of the research is the use of the specified terminology in scientific and regulatory documents.

The subject of the study is the context in which certain elements are used and the dynamics of their use.

Research methods are text mining (text mining), intelligence analysis, comparative analysis, RS-analysis.

The results of this thesis were partially presented at the XVII Scientific and Practical Conference of students, entrants and young specialists "Theoretical and Applied Problems of Physics, Mathematics and Informing"(April 26-27 2019, Kyiv).

CYBERSPACE, CYBERSECURITY, CYBER THREAT, CYBER RISK, TERMINOLOGY, TEXT MINING, RS-ANALYSIS

ЗМІСТ

ВСТУП

Актуальність дослідження. Наразі існує велика кількість джерел, які використовують терміни Кіберпростір (Cyberspace), Кібербезпека (Cybersecurity), Кіберзагрози (Cyber threats), Кібернетичний ризик (Cyber risk). Причому автори, які використовують ці категорії, розуміють їх в різний спосіб, бачення даних категорій інколи суттєво відрізняється. Запровадження нового закону України “Про основні засади забезпечення кібербезпеки України” [3], який дає чіткі визначення понять “кібербезпека”, “кіберзагроза”, “кіберпростір”, спричинило значний поштовх у розвитку відповідної сфери. Швидкість становлення галузі кібербезпеки зумовила певний брак досліджень в напрямку розуміння змісту термінів, що тут виникають, фахівцями та суспільством.

Актуальність та новизна даної роботи зумовлена тим, що незважаючи на достатньо довгий час вживання даних категорій, все ж залишились неоднозначності в трактуванні сфери та особливостей їх застосування, які досі не вирішені у існуючих роботах. До того ж, автоматизованому дослідженню існуючих відкритих корпусів текстів на предмет використання даних термінів досі не було приділено належної уваги.

Метою даної роботи є вирішення термінологічної проблеми у сфері кібербезпеки та чітке розділення інформаційної та кібернетичної сфер.

Об’єкт дослідження – використання зазначеної термінології у наукових та нормативно-правових документах.

Предметом дослідження є контекст в якому використовуються визначені елементи та динаміка обсягу їх використання.

Результати цієї роботи були частково представлені на XVII Науково-практичній конференції студентів, аспірантів та молодих вчених "Теоретичні і прикладні проблеми фізики, математики та інформатики"(26-27 квітня 2019р., м. Київ).

1 АНАЛІЗ СЕМАНТИКИ ТЕРМІНІВ

Основними елементами у сфері кібербезпеки в контексті загрози є:

1. Кіберпростір (Cyberspace)
2. Кібербезпека (Cybersecurity)
3. Кіберзагрози (Cyber threats)
4. Кібернетичний ризик (Cyber risk)

Саме визначення цих термінів є ключовими моментом у вирішенні термінологічної проблеми даної сфери та створення ефективних механізмів протидії кіберзлочинам. Попри те що ці поняття використовуються в законодавчій та професійній практиці різних держав, їх зміст є досі невизначеним що також створює нормативно-правову проблему та ускладнює формування відповідних державних документів, які повинні визначати підходи до проблем кібербезпеки.

1.1 Визначення терміну “Кіберпростір”

Безумовно, ключовим моментом у формування тезаурусу сфери кібербезпеки є визначення поняття *кіберпростір*. Кіберпростір називають одним найбільшим нерегульованих і неконтрольованих доменів в історії людства, який також є унікальном, оскільки він є штучним доменом, створеним людьми з дуже короткою історією.

Кіберпростір є широко поширеною, взаємопов'язаною цифровою технологією. Цей термін увійшов до популярної культури з наукової фантастики та мистецтва, але тепер використовується технологічними стратегами, фахівцями з безпеки, урядовими, військовими та промисловими лідерами та

підприємцями для опису сфери глобального технологічного середовища. Інші вважають, що кіберпростір - це просто умовне середовище, в якому відбувається зв'язок через комп'ютерні мережі. Так наразі визначає кіберпростір Оксфордський словник[4].

Вперше термін «кіберпростір» було введено у вжиток письменником Вільямом Гібсоном у 1982 р. в новелі «Палаючий Хром» («Burning Chrome»). У 1984 році це поняття було більш детально розкрито у творі «Нейромант» («Neuromancer»). На думку Гібсона, кіберпростір (cyberspace) — це злагоджена галюцинація, яку щодня зазнають мільярди звичайних операторів у всьому світі. Це логічне представлення відомостей, збережених в пам'яті та на магнітних носіях комп'ютерів всього розумного людства. Потоки даних, що протікають у просторі розуму; скупчення та сузір'я інформації [5].

Як зможемо побачити далі в цій роботі, слово кіберпростір стало популярним у 1990-х роках, коли використання Інтернету, створення мереж і цифрове спілкування різко зростали, і термін "кіберпростір" міг представляти багато нових ідей і явищ, які виникали. Цьому зокрема присвячена робота “The varieties of cyberspace: Problems in definition and delimitation”[6].

На даний момент не існує спільних визначень кіберпростору на науковому рівні, і кожен уряд використовує окреме визначення. (Достатньо лише розглянути, наприклад, що за Ф. Д. Крамером існує 28 різних визначень терміну кіберпростір[7]).

Деякі, такі як Канада, називають кіберпростір "глобальними общинами", інші, як Німеччина, обмежують визначення всесвіту Інтернету, явно виключаючи інші типи мереж між комп'ютерами. МСЕ (Міжнародний союз електрозв'язку), агентство Організації Об'єднаних Націй у сфері телекомунікацій, визначає кіберпростір як "системи та послуги, пов'язані безпосередньо або опосередковано з Інтернетом, телекомунікаціями та комп'ютерними мережами"[8].

Це визначення здається неповним через відсутність деяких ключових компонентів. Міжнародна організація зі стандартизації (ISO / IEC 27032: 2012) визначає кіберпростір таким чином: "складне середовище, що виникає в результаті взаємодії людей, програмного забезпечення та послуг в Інтернеті за допомогою технологічних пристроїв і мереж, підключених до нього, що не існують у будь-якій фізичній формі"[9]. У цьому визначенні не йдеться про фізичну інфраструктуру, без якої кіберпростір не може існувати.

На мій погляд, більш корисним є вивчення еволюції різних визначень кіберпростору, які використовувалися в Пентагоні. У жовтні 2007 року в «Словнику військових та асоційованих термінів» кіберпростір є «національним середовищем, в якому оцифрована інформація передається через комп'ютерні мережі»[10].

Ця ж публікація, оновлена в серпні 2009 року, визначає кіберпростір наступним чином: "Глобальний домен в інформаційному середовищі, що складається з взаємозалежної мережі інфраструктур інформаційних технологій, включаючи Інтернет, телекомунікаційні мережі, комп'ютерні системи та вбудовані процесори та контролери"[11].

У версія документа, яка була опублікована в червні 2013 року, кіберпростір визначається як: "Глобальний домен в інформаційному середовищі, що складається з взаємозалежної мережі інфраструктур інформаційних технологій і резидентних даних, включаючи Інтернет, телекомунікаційні мережі, комп'ютерні системи та вбудовані процесори і контролери"[12].

Це ж визначення залишається актуальним і наразі. Цікаво відзначити, що протягом кількох років визначення 2007 року було радикально змінено, і як слова "резидентні дані" були додані в останньому визначенні. Ці швидкі зміни в доктрині є відчутним знаком зростаючого сприйняття стратегічної важливості кібер-явища.

Визначення кібернетичного простору, що використовується в Італії, дуже схоже на визначення, яке використовує Міністерство оборони США. У D.P.C.M.

24 січня 2013 року кіберпростір визначається як "загальна система взаємопов'язаних комп'ютерних інфраструктур, включаючи апаратні засоби, програмне забезпечення, дані та користувачів, а також логічні зв'язки між ними, незалежно від того, як вони створені"[13].

Проте, включення терміну "дані", позбавленого будь-яких інших специфікацій, може, призвести до неправильного тлумачення. Тому потрібно ввести відмінність: лише певні типи даних (наприклад, записи DNS, дані для реалізації комунікаційних протоколів тощо) є складовими елементами кіберпростору, тоді як більшість даних, що подорожують та/або проживають у мережах просто містяться в кіберпросторі; іншими словами, вони можуть бути там або не бути там, тоді як без наявності деяких характерних даних кіберпростір не може існувати.

Нижче наведено огляд офіційних визначень кіберпростору в державних документах:

1) “Кіберпростір - це електронний світ, створений взаємопов'язаними мережами інформаційних технологій і інформацією про ці мережі. Це глобальна спільнота, де більше 1,7 мільярда (у версії 2018 року 3 мільярда) людей пов'язані між собою для обміну ідеями, послугами та дружбою”.

—— Канада, “Canada’s Cyber Security Strategy”, 2018. [14]

2) “Кіберпростір - це віртуальний простір усіх IT-систем, пов'язаних на рівні даних у глобальному масштабі.

3) Основою для кіберпростору є Інтернет як універсальна і загальнодоступна мережа зв'язку та транспортування, яка може бути доповнена та розширена будь-якою кількістю додаткових мереж передачі даних. IT-системи в ізольованому віртуальному просторі не є частиною кіберпростору.”

—— Німеччина, “Cyber Security Strategy for Germany”, 2018. [15]

4) “Кіберпростір - це глобальна мережа взаємозалежних інфраструктур інформаційних технологій, телекомунікаційних мереж і систем комп'ютерної обробки, в яких відбувається онлайн-спілкування.”

—— Нова Зеландія, “New Zealand Cyberspace Strategy”, 2015. [16]

5) “Кіберпростір - взаємозалежна мережа інфраструктур інформаційних технологій, що включає в себе Інтернет, телекомунікаційні мережі, комп'ютерні системи, підключені до Інтернету пристрої та вбудовані процесори та контролери. Він може також посилатися на віртуальний світ або домен як досвідчене явище, або абстрактне поняття.”

—— Велика Британія, “NATIONAL CYBER SECURITY STRATEGY 2016-2021”. [17]

6) “Кіберпростір - це складне середовище, що є результатом взаємодії людей, програмного забезпечення та послуг в Інтернеті за допомогою технологічних пристроїв і мереж, пов'язаних з нею, які не існують у жодній фізичній формі”

—— ISO/IEC, ISO/IEC 27032, “Guidelines for cybersecurity (DRAFT)”, 2011. [9]

7) “Кіберпростір - глобальна сфера в інформаційному середовищі, що складається з взаємозалежної мережі інфраструктур інформаційних технологій і резидентних даних, включаючи Інтернет, телекомунікаційні мережі, комп'ютерні системи та вбудовані процесори та контролери.”

—— Сполучені Штати Америки, “DoD Dictionary of Military and Associated Terms”, 2013-2018. [18]

8) “Кіберпростір - це загальна система взаємопов'язаних комп'ютерних інфраструктур, включаючи обладнання, програмне забезпечення, дані та користувачів, а також логічні зв'язки між ними, незалежно від того, як вони встановлені”

—— Італія, “Decree of the President of the Council of Ministers (DPCM)”, 2013. [13]

9) “Кіберпростір - середовище (віртуальний простір), яке надає можливості для здійснення комунікацій та/або реалізації суспільних відносин, утворене в результаті функціонування сумісних (з’єднаних) комунікаційних

систем та забезпечення електронних комунікацій з використанням мережі Інтернет та/або інших глобальних мереж передачі даних.”

—— Україна, “Про основні засади забезпечення кібербезпеки України”, 2018. [3]

1.1.1 Результати

За проведеним аналізом означень кіберпростору різних країн, можна виділити основні структурні елементи, присутні в означеннях:

- взаємозв’язок із мережею Інтернет та/чи іншими глобальними мережами;
- призначення кіберпростору;
- суб’єкти, які беруть участь у комунікаціях із використанням кіберпростору;
- види інформаційних та комунікаційних систем, які беруть участь у формуванні кіберпростору;
- дані, які фігурують в кіберпросторі;
- види та характеристики зв’язків між системами, даними та особами у кіберпросторі.

Отже, це є сукупність основних елементів, які дають змогу однозначно визначити, що таке кіберпростір. Відсутність у визначеннях одного чи декількох із цих елементів дає змогу трактувати їх довільно, що й призводить до різних розумінь терміну “кіберпростір”.

Аналіз означень показує, що деякі із них роблять акцент лише на технологічному боці явища “кіберпростір”, випускаючи із уваги людський фактор та суспільні відносини, які мають місце при безпосередньому використанні кіберпростору

1.2 Визначення терміну “Кібербезпека”

Водночас, з точки зору практичного застосування найбільш значущою проблемою є все ж визначення поняття “Кібербезпека”, яке залежно від закладеного в нього розуміння впливає на формування кола суб'єктів її забезпечення на національному та міждержавному рівні.

Останніми роками ”Кібербезпека” стала широко вживаним терміном. Проте, як і в багатьох модних “жаргонах”, здається, існує дуже мало розуміння того, що термін насправді тягне за собою. Хоч це не викликає труднощі, коли термін використовується в неформальному контексті, він може потенційно викликати значні проблеми в контексті організаційної стратегії, бізнес-цілей або міжнародних угод.

Нова термінологія почала набувати все більшої популярності з використанням терміну “Cyber Security”. Вона використовувалася протягом попередніх років, але її популярність значно зросла, коли президент США Барак Обама в 2009 році році закликав звернути увагу до кібербезпеки як складової національної безпеки - “I call upon the people of the United States to recognize the importance of cybersecurity and to observe this month with appropriate activities, events, and trainings to enhance our national security and resilience.” [19]

Безпосередній вплив цього прес-релізу на термінологію можна проілюструвати за допомогою пошукових тенденцій Google, які за цей період помітно збільшилися (Рисунок 1.1). Рядки тренду на діаграмі показують загальну кількість пошуків за терміном по відношенню до загальної кількості пошуків, здійснених на Google протягом часу. Ми бачимо стійке зниження пошукових термінів «Комп'ютерна безпека» та «Інформаційна безпека» порівняно з варіантами «кібербезпеки».

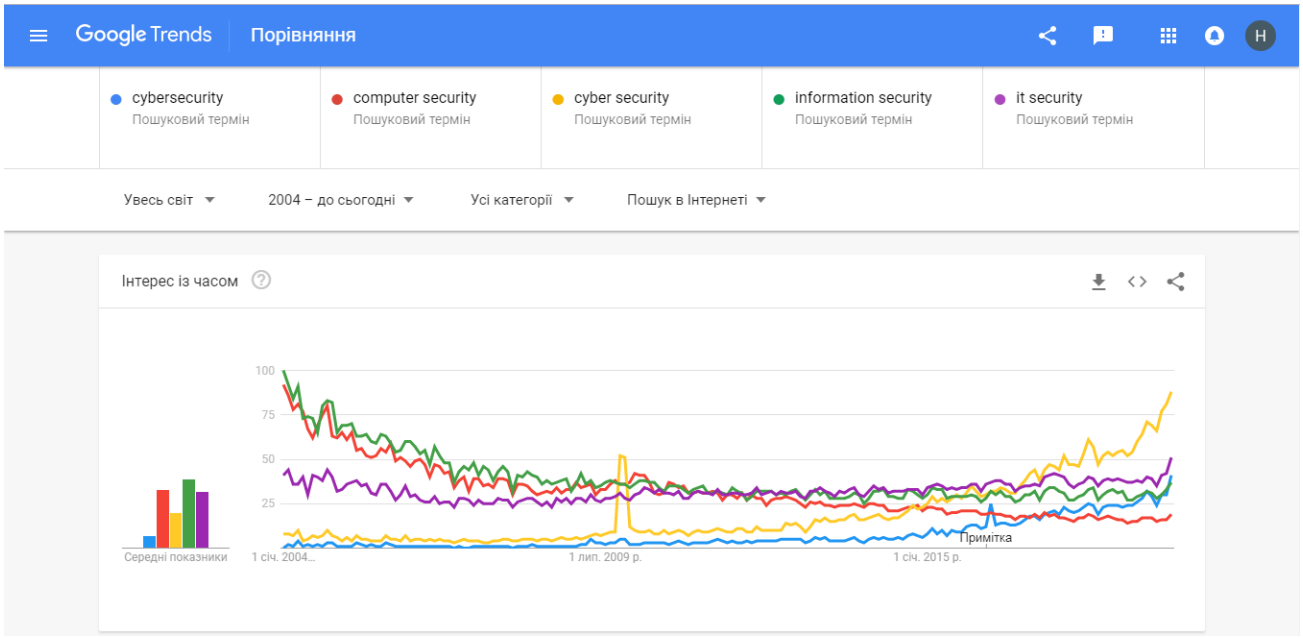

Рисунок 1.1 – Використання терміну “Cyber security” та інших споріднених із ним

У наукових публікаціях термін вперше використано в книзі “Computers & Security” автор Helen Meyer [20]. Визначення терміну не надається. Цитата - “Cypherwiz cracks cybersecurity.”

Оксфордський словник зокрема визначає “cybersecurity” як: “Стан захищеності від кримінального або несанкціонованого використання електронних даних, або заходів, вжитих для досягнення цього”[21].

Огляд термінології які застосовуються у стратегіях забезпечення безпеки в кіберпросторі різних країн:

1) “Кібербезпека: захист цифрової інформації, а також цілісність інфраструктури та передача цифрової інформації. Зокрема, кібербезпека включає в себе цілий ряд технологій, процесів, практик та заходів реагування та пом'якшення, призначених для захисту мереж, комп'ютерів, програм і даних від нападу, пошкодження або несанкціонованого доступу з метою забезпечення конфіденційності, цілісності та доступності.”

—— Канада, “Canada’s Cyber Security Strategy”, 2018. [14]

2) Німецька доктрина розділяє кібербезпеку на три категорії залежно від об’єктів захисту:

“(Глобальна) кібербезпека є бажаною метою ситуації безпеки ІТ, в якій ризики глобального кіберпростору зводяться до прийнятного мінімуму.

Отже, кібербезпека в Німеччині є бажаною метою ситуації в сфері безпеки ІТ, в якій ризики німецького кіберпростору зводяться до прийнятного мінімуму. Кібербезпека (у Німеччині) - це сума відповідних заходів.

Цивільна кібербезпека фокусується на всіх ІТ-системах для цивільного використання в німецькому кіберпросторі. Військова кібербезпека зосереджена на всіх ІТ-системах для військового використання в німецькому кіберпросторі.”

—— Німеччина, “Cyber Security Strategy for Germany”, 2018 [15]

3) “Кібербезпека - захист систем, підключених до Інтернету (включаючи апаратні засоби, програмне забезпечення та пов'язану з ними інфраструктуру), дані про них та послуги, які вони надають, від несанкціонованого доступу, шкоди або зловживання. Це включає в себе шкоду, завдану навмисно оператором системи, або випадково, внаслідок невиконання процедур безпеки або маніпуляції з нею.”

—— Велика Британія, “NATIONAL CYBER SECURITY STRATEGY 2016-2021”. [17]

4) “Безпека кіберпростору - дії, вжиті в захищеному кіберпросторі для запобігання несанкціонованому доступу, експлуатації або пошкодженню комп'ютерів, систем електронних комунікацій та інших інформаційних технологій, включаючи інформаційні технології платформи, а також інформацію, що міститься в ньому, для забезпечення її доступності, цілісності , аутентифікації, конфіденційності та невідмовності.”

—— Сполучені Штати Америки, “DoD Dictionary of Military and Associated Terms”, 2013-2018. [18]

5) “Кібербезпека - збереження конфіденційності, цілісності та доступності інформації в кіберпросторі”

—— **ISO/IEC, ISO/IEC 27032, “Guidelines for cybersecurity (DRAFT)”, 2011.** [9]

6) “Кібербезпека - захищеність життєво важливих інтересів людини і громадянина, суспільства та держави під час використання кіберпростору, за якої забезпечуються сталий розвиток інформаційного суспільства та цифрового комунікативного середовища, своєчасне виявлення, запобігання і нейтралізація реальних і потенційних загроз національній безпеці України у кіберпросторі.”

—— **Україна, “Про основні засади забезпечення кібербезпеки України”, 2018.** [3]

1.2.1 Результати

Визначення кібербезпеки даються в контексті «стану захищеності», «захисту» (дій стосовно забезпечення захищеності) тощо. Також відзначаються: складові кібербезпеки (технології, процеси, практики), об'єкти кібербезпеки (мережі, комп'ютери, програми, дані), мета кібербезпеки, і основні загрози, яким вона має протистояти (наприклад, несанкціонований доступ, пошкодження або зловживання). В деяких країнах виділяються різні категорії кібербезпеки, в залежності від об’єктів захисту.

Вигідно відрізняється від існуючих визначення, використане в законі України, оскільки воно єдине конкретизує, що об’єктом захисту є захищеність життєво важливих інтересів людини і громадянина, суспільства та держави під час використання кіберпростору, акцентуючи увагу на людському та суспільному факторі.

1.3 Визначення терміну “Кіберзагроза”

Згідно даних пошукових систем, У наукових публікацях термін вперше використано у книзі “*Inside the Army Vol*” автор Jason Sherman. Цитата, в якій використано словосполучення: “*DOD RETHINKING **CYBER THREATS***” [22]

Оксфордський словник надає такие визначення “cyberthreat”: “Можливість зловмисної спроби пошкодити або зірвати комп'ютерну мережу або систему.” [23]

Порівняно із визначеннями «кіберпростору» та «кібербезпеки», визначення «кіберзагрози» є достатньо однотипними для різних країн. У визначенні терміну зазначається про ті явища, та чинники, які завдають шкоди та створюють небезпеку для об'єктів захисту. (напр., **Канада** -“Кібер загроза: Актор загрози “*A threat actor*”, що використовує Інтернет, який використовує перевагу знання уразливість в продукті для цілей експлуатації мережі та інформації, яку мережа зберігає.”**, “Canada’s Cyber Security Strategy”, 2018.** [14]**)**

Цікавим є те, що в західних країнах в якості об'єктів захисту постає апаратне, програмне забезпечення, інфраструктура (напр., **Велика Британія** - “Кібер-загроза - все, що здатне поставити під загрозу безпеку або завдати шкоди інформаційним системам та пристроям, підключених до Інтернету (включаючи апаратне забезпечення, програмне забезпечення та пов'язану з ним інфраструктуру), дані про них та послуги, які вони надають, насамперед за допомогою кібер-засобів.” **“NATIONAL CYBER SECURITY STRATEGY 2016-2021”.** [17]), а в українському законодавстві надано більш широке визначення, яке стосується небезпек національним інтересам України, та стану кібербезпеки кіберзахисту об’єктів різного характеру. (“Кіберзагроза - наявні та потенційно можливі явища і чинники, що створюють небезпеку життєво важливим національним інтересам України у кіберпросторі, справляють

негативний вплив на стан кібербезпеки держави, кібербезпеку та кіберзахист її об’єктів” **“Про основні засади забезпечення кібербезпеки України”, 2018.** [3])

1.4 Визначення терміну “Кібернетичний ризик”

За результатами пошуку, у наукових публікацях термін cyber risk вперше використано в статті “Risks of Cyber Attack to Water Utility Supervisory Control and Data Acquisition Systems” автори Barry C. Ezell, Yacov Y. Haimes and James H. Lambert. Цитата: *“Using system decom- position, sources of* ***cyber risk*** *through SCADA are identified.”* [24]

Хоча ще в публікації 1995 року у жуналі *Michigan Law Review* використовувався вираз: “*The freedom of speech at* ***risk in cyberspace****: obscenity doctrine and a frightened university's censorship of sex on the Internet.*” [25]

В законодавстві України та інших західних держав безпосередньо визначення кібернетичного ризику не надається

Висновки до розділу 1

У значній мірі опираючись на роботи Качинського А.Б.[1] у цій сфері та зважаючи на аналіз термінології наведений вище запропоновані наступні визначення:

Кіберпростір - це середовище, яке надає можливості для здійснення комунікації та управління особистих, суспільних та державних відносин, утворене в результаті функціонування сумісних комунікаційних систем та забезпечення електронних комунікацій з використанням мережі Інтернет та/або інших глобальних мереж передачі даних.

Кібернетична безпека (кібербезпека) – це стан захищеності життєво важливих інтересів особи, суспільства та держави від зовнішніх і внутрішніх загроз, джерелом яких є кібернетичний простір.

Кібернетичні загрози (кіберзагрози) – це прогнозовані, але не контрольовані явища, події та процеси, що відбуваються в кібернетичному просторі, і можуть завдати значних збитків матеріальним і духовним цінностям особи, суспільства та держави

Кібернетичний ризик (кіберризик) – це прогнозована векторна величина збитку, що пов’язана із реалізацією кібернетичної загрози, і дорівнює добутку ймовірності її реалізації на ймовірність величини можливого збитку від даної загрози.

2 ДОСЛІДЖЕННЯ ДИНАМІКИ ВЖИВАНЬ ТЕРМІНІВ

2.1 Дослідження наукових публікацій

У ході роботи було досліджено зміну кількості використань наведених вище термінів в наукових публікаціях з часом. Для цього використовували наступні три сервіси: Google Scholar, JSTOR, ScienceDirect як одні із найбільших ефективних онлайнових сервісів, що надають доступ до наукових публікацій.

2.1.1 Опис сервісів

Google Scholar[26] — вільна доступна пошукова система, яка індексує повний текст наукових публікацій всіх форматів і дисциплін. Індекс Google Scholar включає в себе більшість рецензованих онлайн-журналів Європи та Америки найбільших наукових видавництв. Незважаючи на те, що Google не публікує розмір бази даних Google Scholar, дослідники оцінили, що він містить приблизно 389 мільйонів документів, включаючи статті, цитати та патенти, що робить його найбільшим у світі науковим пошуковим центром у січні 2018 року[27]. Більш ранні статистичні оцінки, опубліковані в PLOS ONE з використанням методу Mark і recapture, оцінювали охоплення в приблизно 80–90% всіх статей, опублікованих англійською мовою, з розміром в 100 мільйонів[28].

JSTOR[29] (Скорочення від англ. *Journal STORage*) — цифрова повнотекстова база даних наукових журналів (на різних європейських мовах), а також книг (гуманітарні науки, тільки англійською мовою). JSTOR надає доступ до більш ніж 12 мільйонів наукових журналів, книг та первинних джерел у 75 дисциплінах.

ScienceDirect[30] — одна з найбільших онлайн колекцій опублікованих наукових досліджень. Належить голландському видавництву Elsevier. На лютий 2016 року містила майже 12 мільйонів елементів контенту з більш ніж 3500 журналів і понад 34000 електронних книг, довідників, наукових збірників. Статті згруповано в чотири основні розділи: фізичні і технічні науки, природничі науки, медичні науки та соціальні і гуманітарні науки. Анотації більшості статей знаходяться у вільному доступі. Доступ до повного тексту статті вимагає передплати.

2.1.2 Результати

Англомовні джерела досліджувались за допомогою всіх трьох сервісів. Серед ще не до кінця встановлених варіантів правопису(наприклад “Cyber threat”,” Cyberthreat”, ”Cyber-threat”) використовувались найбільш часто вживані.

Україномовні джерела досліджувались тільки за допомогою Google Scholar, оскільки він єдиний містить базу україномовних публікацій. В результаті було складено таблиці вживання термінів по роках за різними ресурсами за 1982-2018 роки.

Таблиця 2.1 – Кількість публікацій, які використовують термін “Cyberspace” по рокам

Рік	Google Scholar	JSTOR	ScienceDirect
2018	20700	478	409
2017	22000	734	377
2016	22500	798	338
2015	23000	814	408
2014	23300	1141	384

Продовження таблиці 2.1

Рік	Google Scholar	JSTOR	ScienceDirect
2013	24000	1203	378
2012	23100	1455	284
2011	20900	1310	296
2010	20100	1201	248
2009	18300	1354	255
2008	17900	1149	227
2007	16400	1048	246
2006	14600	954	211
2005	14300	966	225
2004	14000	959	232
2003	12400	993	239
2002	13100	1057	272
2001	12300	1195	261
2000	11300	1142	284
1999	9170	1177	271
1998	7670	925	255
1997	6880	868	366
1996	5990	730	341
1995	3040	468	242
1994	1740	239	108
1993	825	96	52
1992	385	61	21
1991	341	32	14
1990	143	21	3
1989	66	5	0
1988	55	6	1
1987	17	0	0
1986	46	1	0
1985	25	1	0
1984	27	0	0
1983	13	0	0
1982	19	0	0

Таблиця 2.2 - Кількість публікацій, які використовують термін “Cybersecurity” по рокам

Рік	Google Scholar	JSTOR	ScienceDirect
2018	16100	233	753
2017	17000	436	433
2016	15200	451	321
2015	12100	346	239
2014	7620	603	143
2013	5900	469	158
2012	4470	379	49
2011	3910	372	76
2010	3500	237	32
2009	2710	159	43
2008	1530	86	25
2007	1160	30	39
2006	1120	47	17
2005	1260	40	20
2004	962	61	71
2003	906	51	16
2002	792	19	19
2001	589	5	7
2000	527	5	1
1999	439	1	2
1998	432	1	0
1997	178	1	0
1996	204	0	3

Таблиця 2.3 - Кількість публікацій, які використовують біграму “Cyber threats” по рокам

Рік	Google Scholar	JSTOR	ScienceDirect
2018	4510	71	203
2017	4130	116	158
2016	3630	169	145
2015	2690	213	159
2014	2440	172	100
2013	2050	193	107
2012	1480	169	32
2011	1120	138	44
2010	849	72	23
2009	623	82	18
2008	449	30	8
2007	312	11	28
2006	285	9	9
2005	281	5	8
2004	200	10	10
2003	176	10	6
2002	147	18	6
2001	135	27	11
2000	80	13	3
1999	46	2	5
1998	50	4	4
1997	22	0	7
1996	21	4	2

Таблиця 2.4 - Кількість публікацій, які використовують біграму “Cyber risk” по рокам

cyber risk	Google Scholar	JSTOR	ScienceDirect
2018	1270	16	87
2017	1270	31	73
2016	1030	27	55
2015	743	27	38
2014	530	26	23
2013	358	23	31
2012	217	7	10
2011	164	7	6
2010	151	9	9
2009	102	4	4
2008	108	0	3
2007	83	1	3
2006	68	4	2
2005	66	9	4
2004	33	5	3
2003	39	11	1
2002	23	1	4
2001	24	1	1
2000	11	0	0
1999	2	0	0
1998	6	0	0
1997	0	0	0
1996	1	0	0
1995	2	0	0

Таблиця 2.5 - Кількість україномовних публікацій наявних у Google Scholar

Рік	Кіберпростір	Кібербезпека	Кіберзагрози	Кібер ризик
2018	175	317	52	5
2017	260	476	74	2
2016	243	229	40	2
2015	176	76	27	4
2014	143	65	27	0
2013	189	63	19	0
2012	117	36	6	0
2011	112	29	5	0
2010	70	7	4	0
2009	50	7	1	0
2008	31	8	1	0
2007	12	9	0	0
2006	17	37	0	0
2005	10	1	0	0
2004	8	1	0	0
2003	4	0	0	0
2002	0	1	0	0
2001	2	2	1	0
2000	4	0	0	0
1999	1	0	0	0

2.2 Розвідковий аналіз даних

Для виявлення загальних тенденцій на основі результатів наведених в попередньому пункті побудовано наступі графіки (рисунки 2.1-2.5). Оскільки розміри баз наукових робіт у сервісів значно відрізняються графіки відображають % вживань терміну у конкретний рік порівняно з загальною кількістю використання у конкретному сервісі.

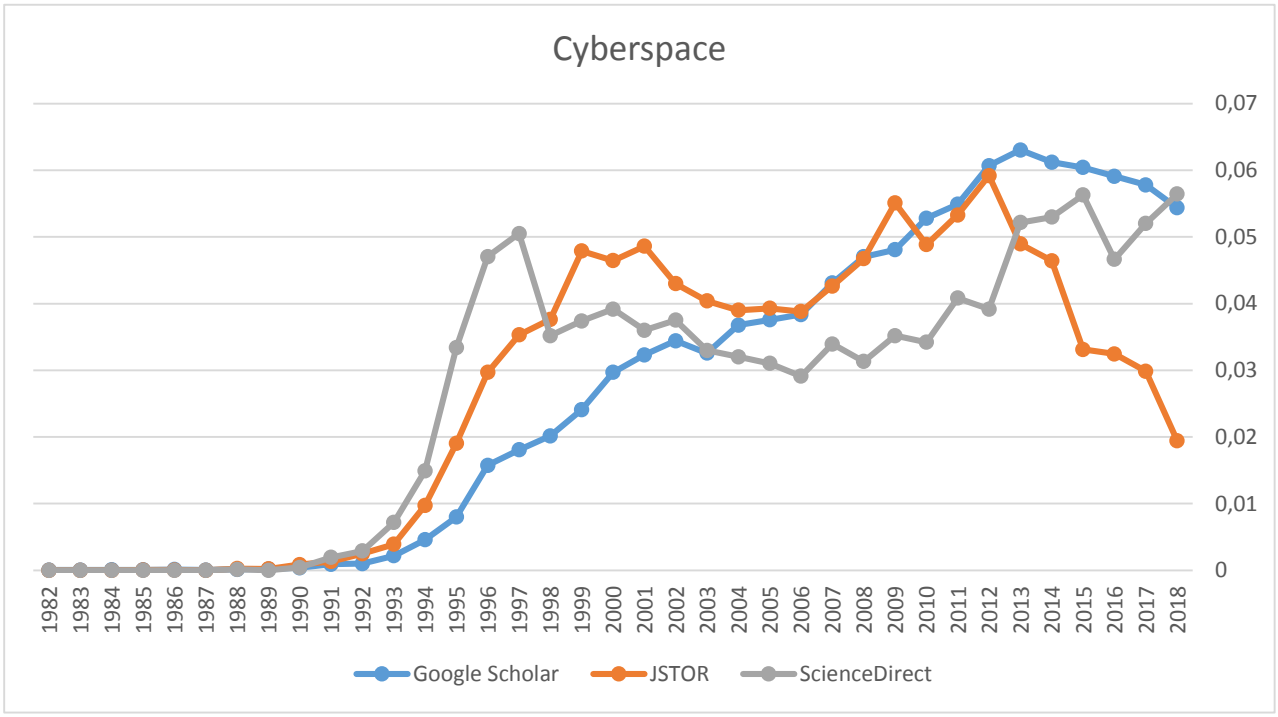

Рисунок 2.1 – Динаміка використання терміну “Кіберпостір” у наукових джерелах

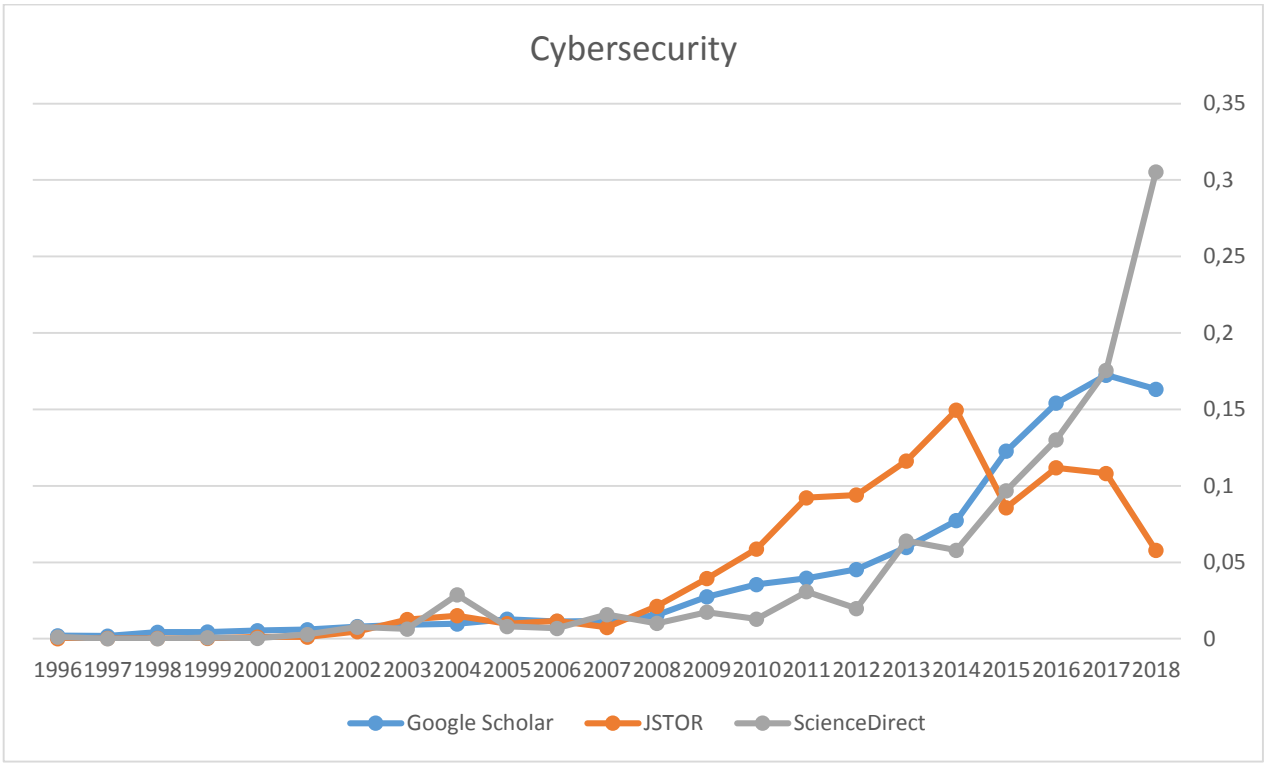

Рисунок 2.2 - Динаміка використання терміну “Кібербезпека” у наукових джерелах

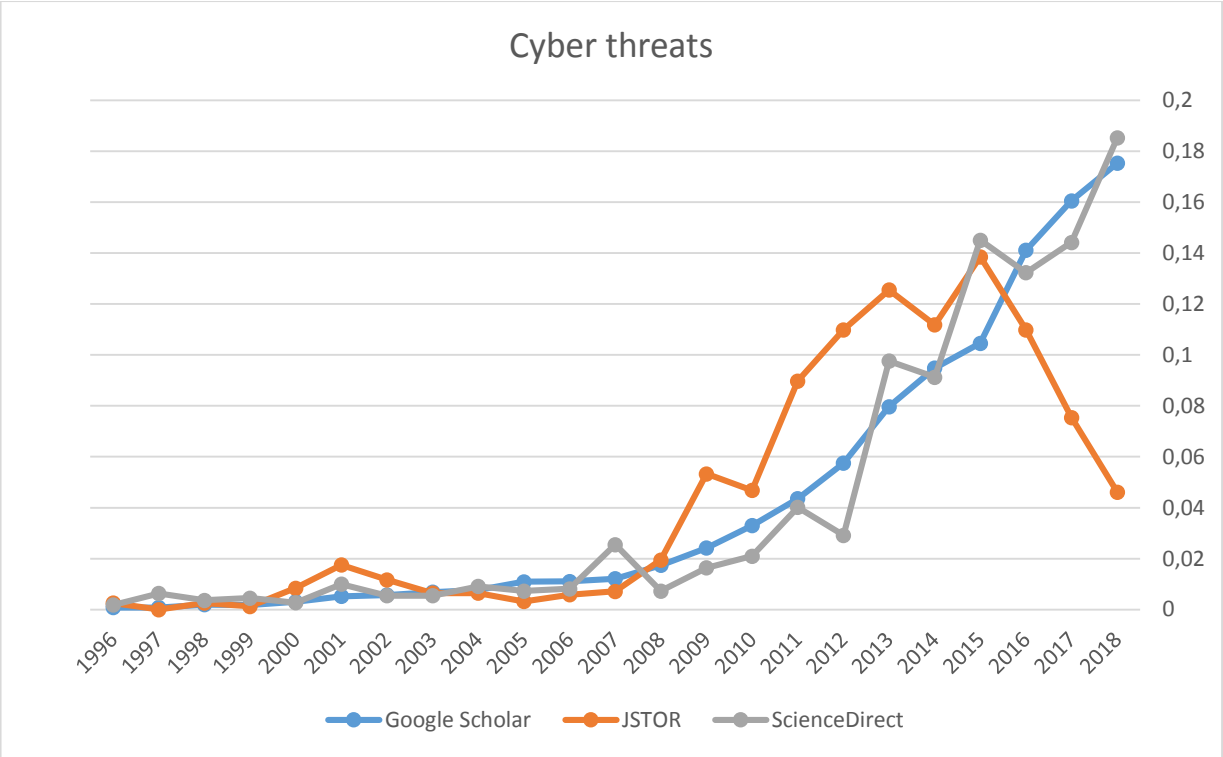

Рисунок 2.3 - Динаміка використання біграми “Кібер загрози” у наукових джерелах

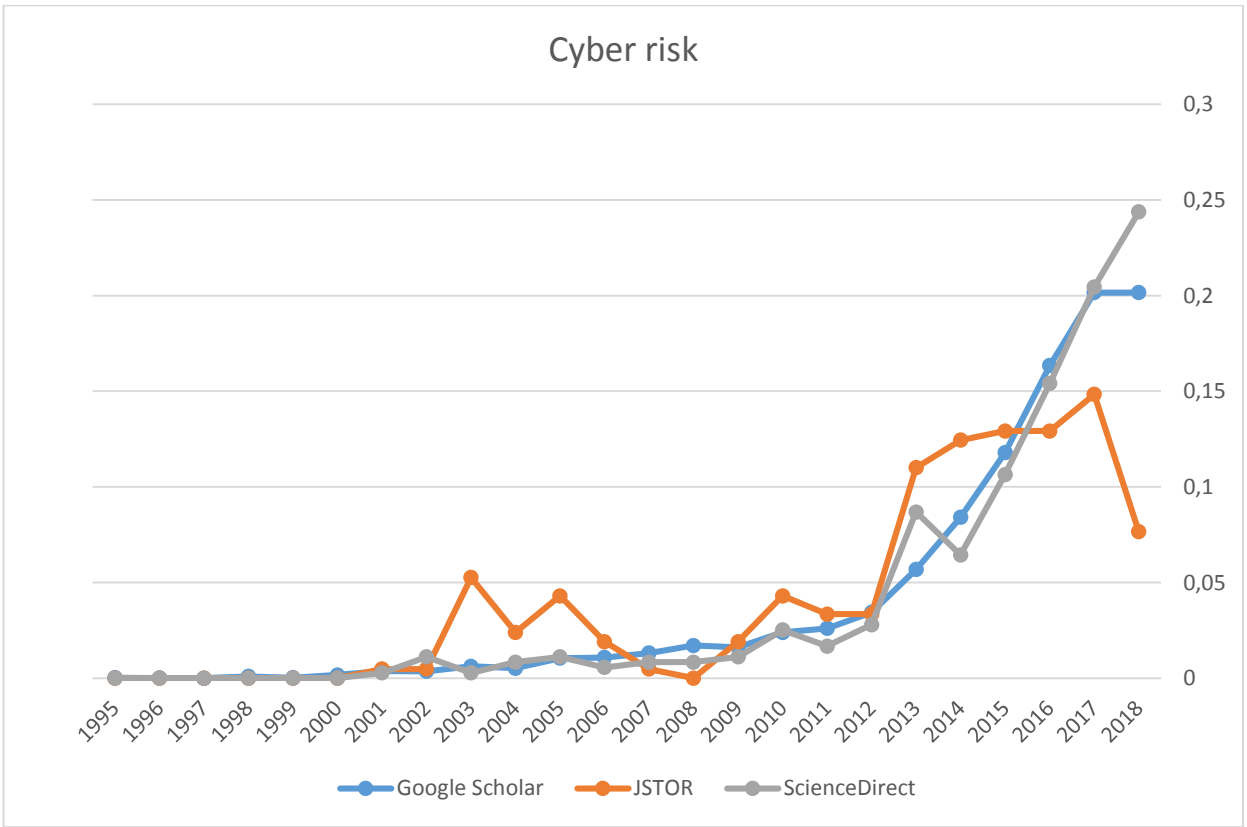

Рисунок 2.4 - Динаміка використання біграми “Кібер ризик” у наукових джерелах

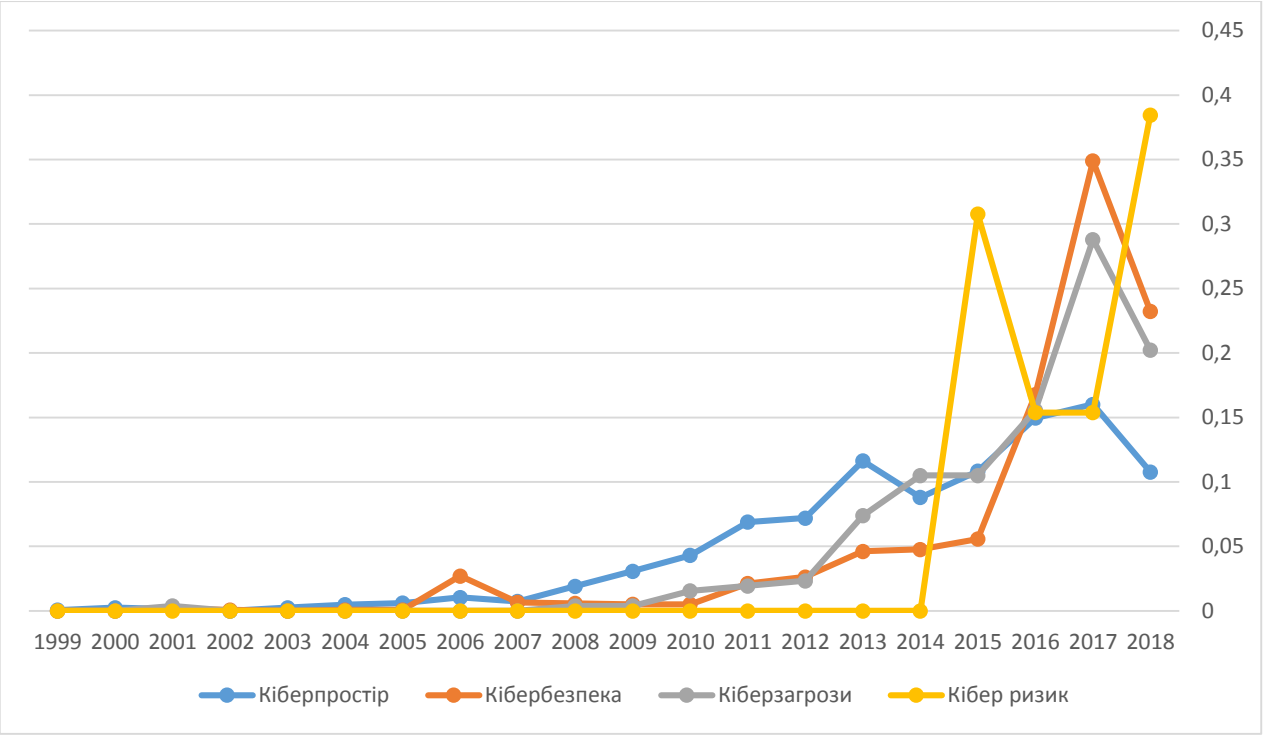

Рисунок 2.5 - Динаміка використання термінології в україномовних наукових джерелах

Як можемо бачити на рисунку 2.1 термін “Cyberspace” починає активно використовуватись з 1995 року і досягає свого піку цитування в 2012-2013 роках, надалі ми можемо констатувати незначне зменшення активності вживання терміну, хоча це може бути пов’язаним з тим що частина нових публікацій досі є платними і поки що не присутня у відкритих базах. Цей момент розглянуто далі у роботі.

Терміни “Cybersecurity”(рисунок 2.2) та “Cyber threats”(рисунок 2.3) починають активно вживатись з 2010 року, “Cyber risk”(рисунок 2.4) на два роки пізніше. Наразі кількість використань усіх термінів стрімко зростає. Зокрема аналізуючи отримані з Google Scholar результати як найбільш точного(через великий обсяг вибірки) можемо стверджувати що цитованість терміну “Cyberspace” зростає лінійно а кількість вживань трьох інших категорій зростає за експоненційним законом.

В українськомовній науковій літературі(рисунок 2.5) дана термінологія почала використовуватись значно пізніше. Зокрема “*Кіберпростір*” відносно активно починає вживатись з 2008 року, “*Кібербезпека*” та “*Кіберзагроза*” у 2010-2011 роках. Термін “*Кібер ризик*” досі має лише декілька згадок у наукових роботах. Хоча у випадку з українськомовною літературою потібно відзначити досить малу кількість публікацій наявних у відкритих онлайн базах.

Також починаючи з 2014 року, активність вживань зазначених термінів різко зростає, це можна пов’язати з фактором російської агресії проти України і виведенням загроз пов’язаних з кіберпростором в порядок денний.

2.3 Аналіз популярності термінів

Попередній пункт не може слугувати для аналізу саме популярності термінології, оскільки потрібно враховувати фактор того що кількість наукових публікацій в цілому з кожним роком зростає. Зокрема у 2018 році порівняно з 2017 кількість наукових робіт в цілому зросла на 5%[31]. Тому для подальшого аналізу було зібрано данні по загальній кількості публікацій на кожному з сервісів у конкретні роки.

У таблиці 2.6 наведено дані по кількості наявних у даний момент наукових робіт за конкретний рік у кожному з використовуваних сервісів. Дані включають в себе роботи всіх напрямів та будь-якими мовами.

Таблиця 2.6 – Кількість наукових робіт доступних на сервісах в цілому по роках

Рік	Google Scholar	JSTOR	ScienceDirect
2018	4990000	64620	737756
2017	3890000	81470	715373
2016	4440000	101778	692321
2015	3560000	181778	672040

Продовження таблиці 2.6

Рік	Google Scholar	JSTOR	ScienceDirect
2014	2650000	182512	636108
2013	3950000	236600	607415
2012	4160000	228039	573953
2011	4260000	233186	543156
2010	4280000	226968	510407
2009	4180000	231083	522614
2008	3990000	225606	489239
2007	3630000	220650	475790
2006	3960000	215901	454134
2005	3470000	214559	415045
2004	3250000	209963	399130
2003	3340000	206985	378686
2002	2900000	205055	330008
2001	3050000	202579	334514
2000	2970000	200490	316964
1999	2450000	199889	294127
1998	2340000	198149	311829
1997	2160000	197149	321602
1996	2310000	195238	307916
1995	2160000	194054	267838
1994	2030000	192308	267838
1993	1850000	191066	258111
1992	1770000	185439	253515
1991	1660000	182743	248832
1990	1660000	179263	232859
1989	1440000	176821	219970
1988	1170000	173538	205960
1987	974000	170125	198380
1986	880000	166814	185835
1985	829000	164839	180307
1984	790000	161999	171522
1983	965000	159637	165850
1982	884000	158296	153369
1981	941000	153651	148449
1980	1090000	153891	141776

На основі даних наведених у таблиці 2.6 за допомогою методу найменших квадратів побудовано графіки регресії (рисунки 2.6-2.7,2.9) та побудований прогноз подальшої динаміки кількості наукових публікацій і цілому до 2023 року.

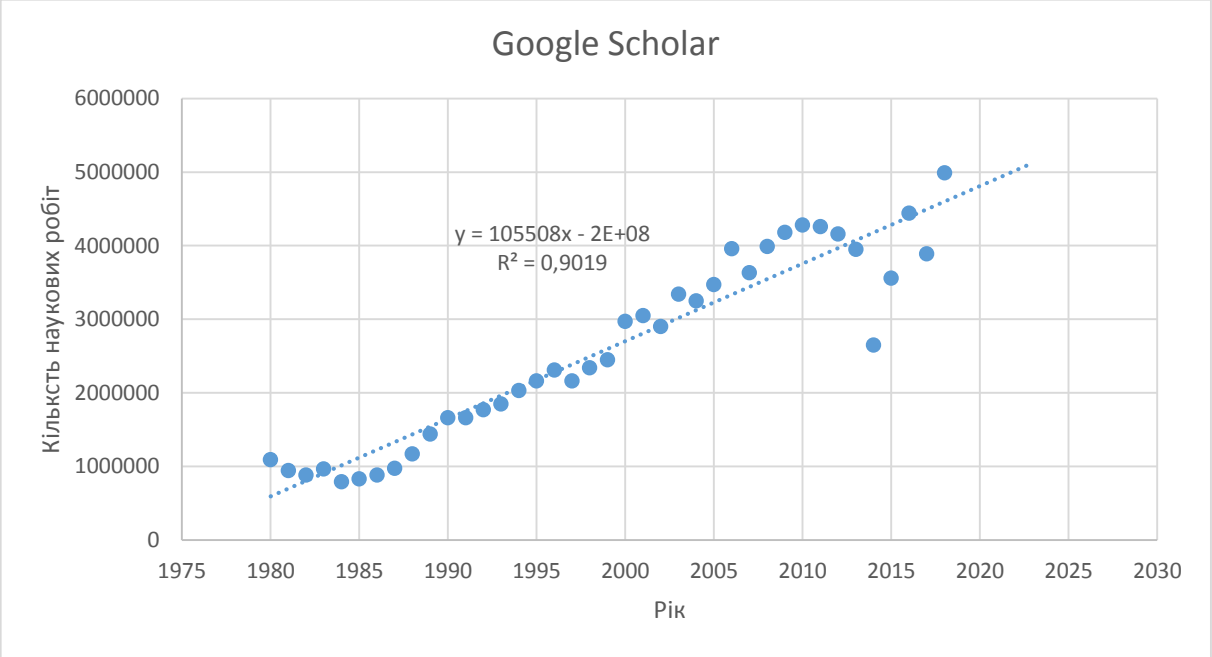

Рисунок 2.6 – Динаміка зростання кількості наукових робіт доступних сервісі Google Scholar

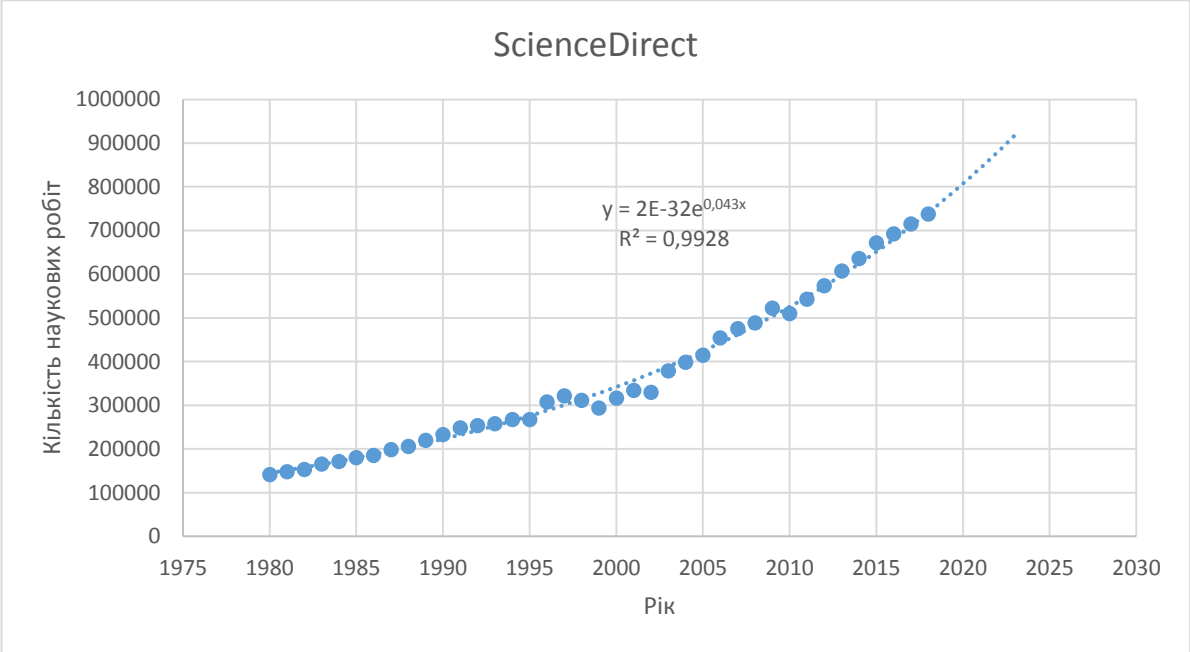

Рисунок 2.7 – Динаміка зростання кількості наукових робіт у доступних сервісі ScienceDirect

По сервісу JSTOR можемо спостерігати 5 викидів які припадають на 2014-2018 роки (рисунок 2.8). Це пов’язано з тим що багато робіт досі є платниими та недоступними для відкритого доступу. Як можемо бачити вище сервіс ScienceDirect цієї проблеми немає а у Google Scholar вона не настільки виражена.

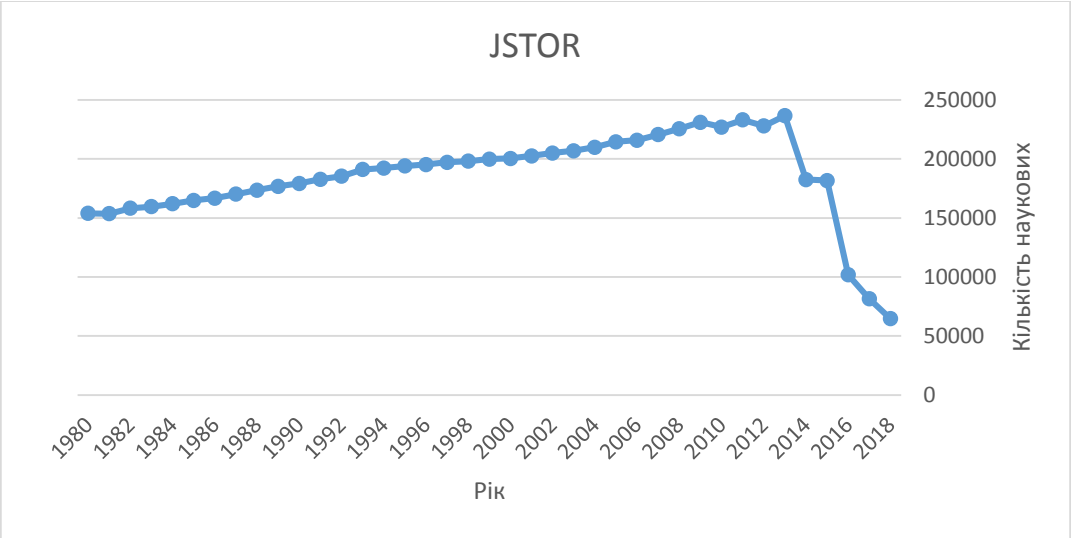

Рисунок 2.8 – Динаміка зростання кількості наукових робіт у сервісі JSTOR

Тому для побудови регресійної моделі по JSTOR(рис 2.9) дані за 2014-2018 не враховувались.

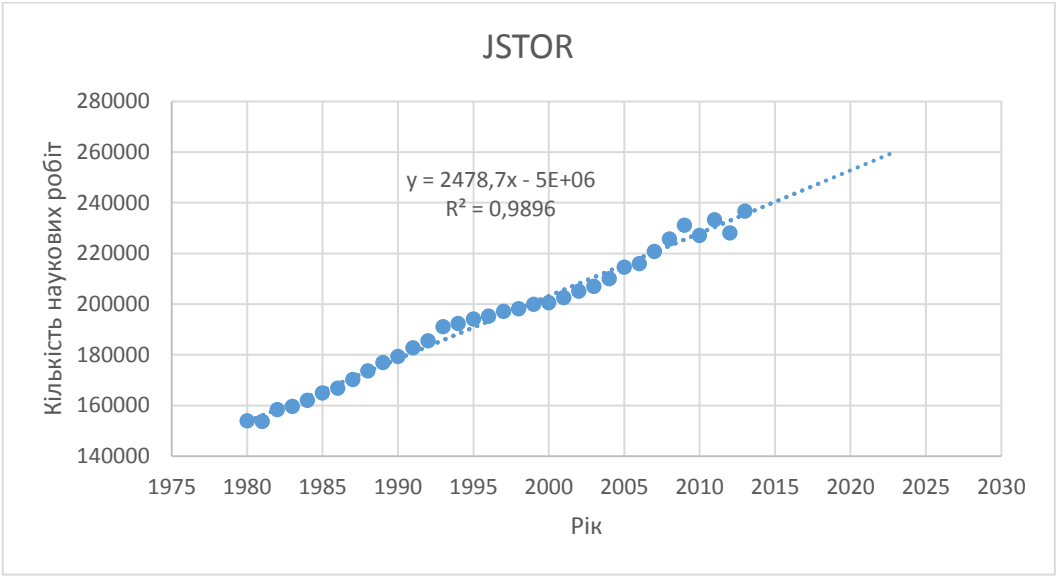

Рисунок 2.9 – Динаміка зростання кількості наукових робіт доступних у сервісі JSTOR

Отже як можемо спостерігати обсяг наукової роботи щороку збільшується тому для перевірки того чи значне зростання кількості цитувань аналізованої термінології не є лише наслідком цього процесу для усіх категорій та використанних сервісів обраховано кореляцію за формулою:

$$r_{xy} = \frac{\sum(x_i - x_{\text{сер}})(y_i - y_{\text{сер}})}{\sqrt{\sum(x_i - x_{\text{сер}})^2 * \sum(y_i - y_{\text{сер}})^2}}$$

Кореляцію обраховували по даним з 1996 по 2018 роки. Кінцеві результати наведені в таблиці 2.7.

Таблиця 2.7 – коефіцієнти кореляції кількості наукових робіт у конкретні роки з вживанням розглянутих елементів та загальної кількості наукових робіт.

	Google Scholar	JSTOR	ScienceDirect
Cyberspace	0,758518335	0,7609725	0,578762221
Cybersecurity	0,55645332	-0,31556	0,777920594
Cyber threats	0,575130019	-0,199329	0,886471013
Cyber risk	0,528506453	-0,570461	0,831730397

Як можемо бачити місцями присутній сильний взаємозвязок, зоркема для севісу ScienceDirect та термінів “Cyber risk” і “Cyber threats” кореляція більша за 0.8 що свідчить про значну залежніть між двома процесами. Тому для у подальшій роботі більшу перевагу надається сервісам Google Scholar та JSTOR у яких для всіх термінів крім “Cyberspace” зв'язок слабкий або відсутній.

В цілому кореляція не є повною. Тому можемо стверджувати що в причинах зростання кількості вживань даної термінології присутні інші фактори окрім загального зростання обсягів наукової роботи.

Для аналізу популярності розглянутої термінології наведених на рисунках 2.10-2.13 зображені графіки відношення кількості наукових робіт з вживанням

шуканих термінів у конкретний рік, до кількості наукових публікацій за цей рік в цілому.

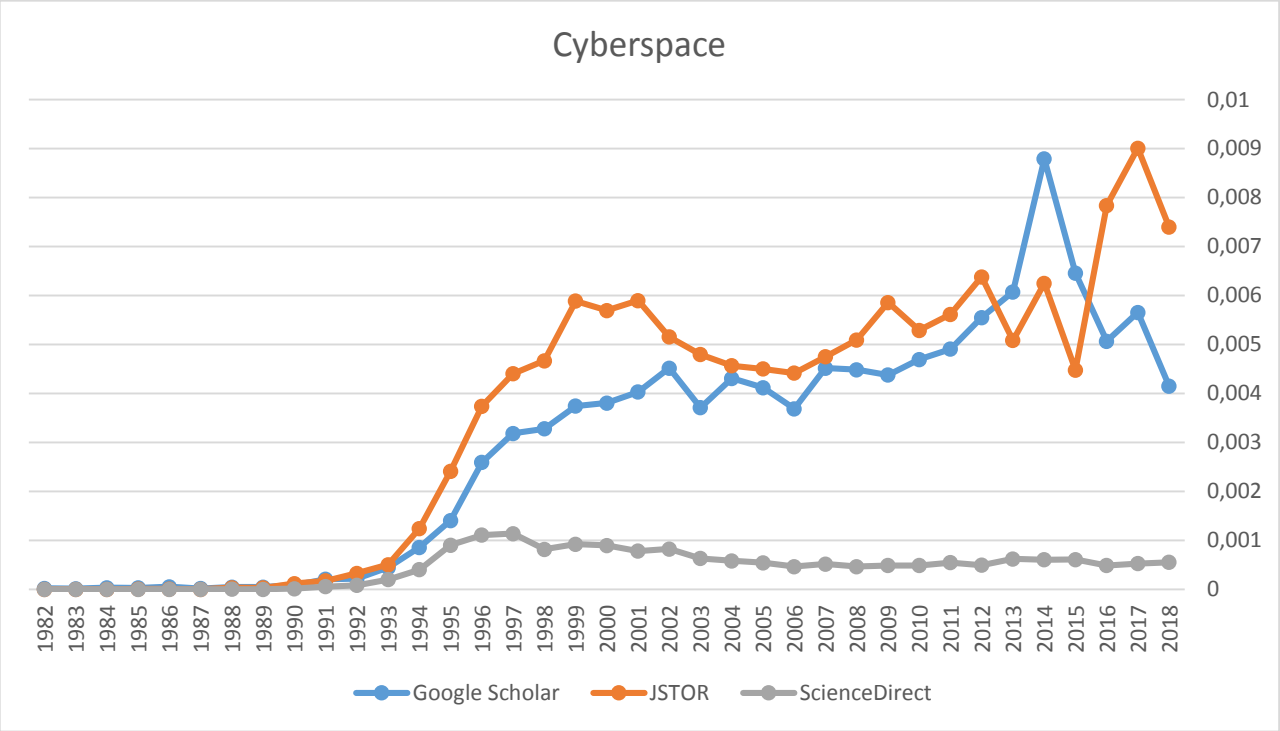

Рисунок 2.10 – Динаміка популярності терміну “Cyberspace”

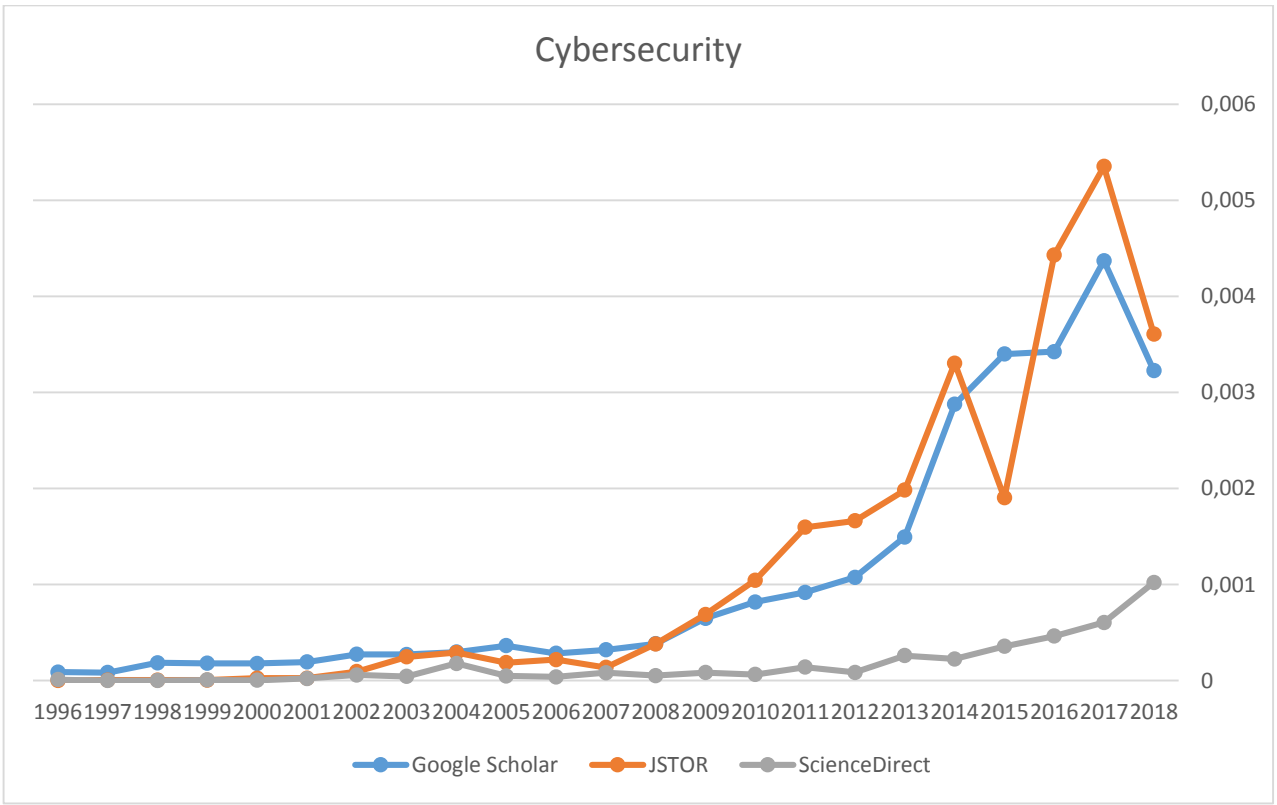

Рисунок 2.11 – Динаміка популярності терміну “Cybersecurity”

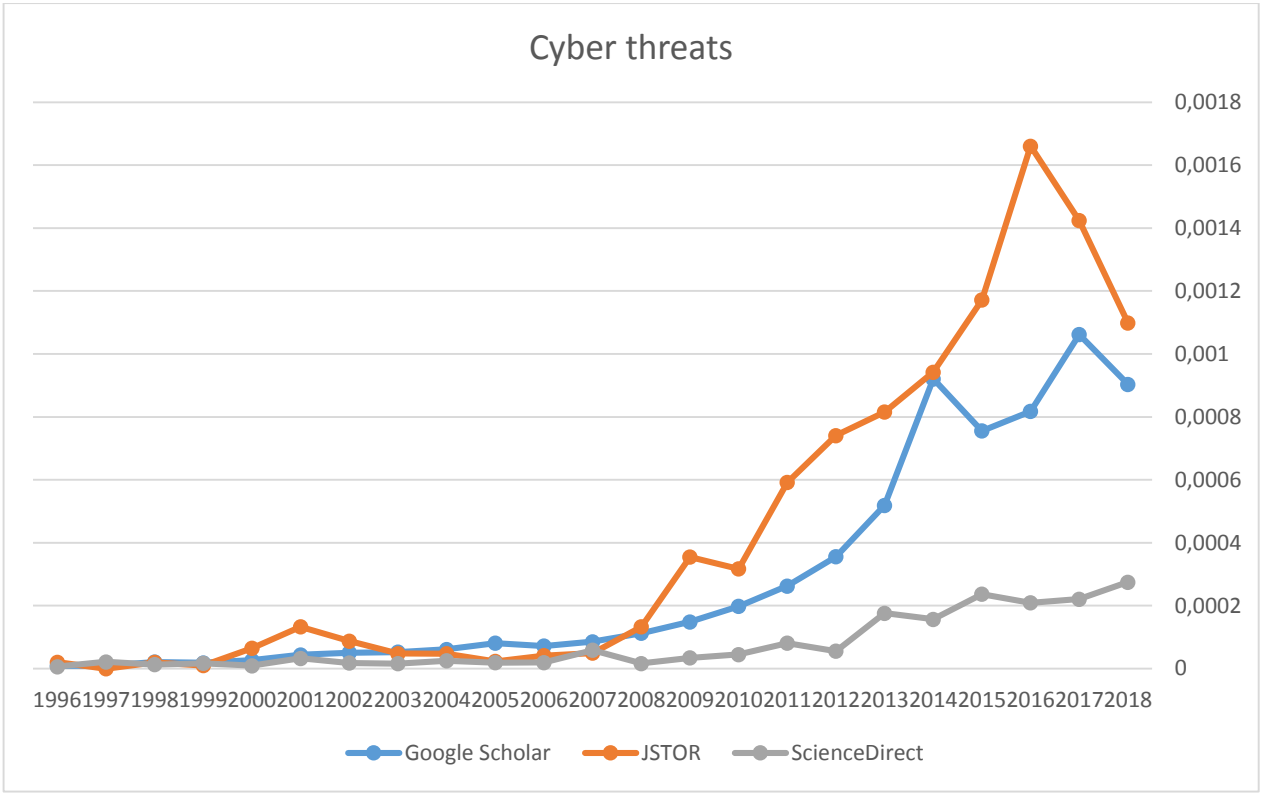

Рисунок 2.12 – Динаміка популярності біграми “Cyber threats”

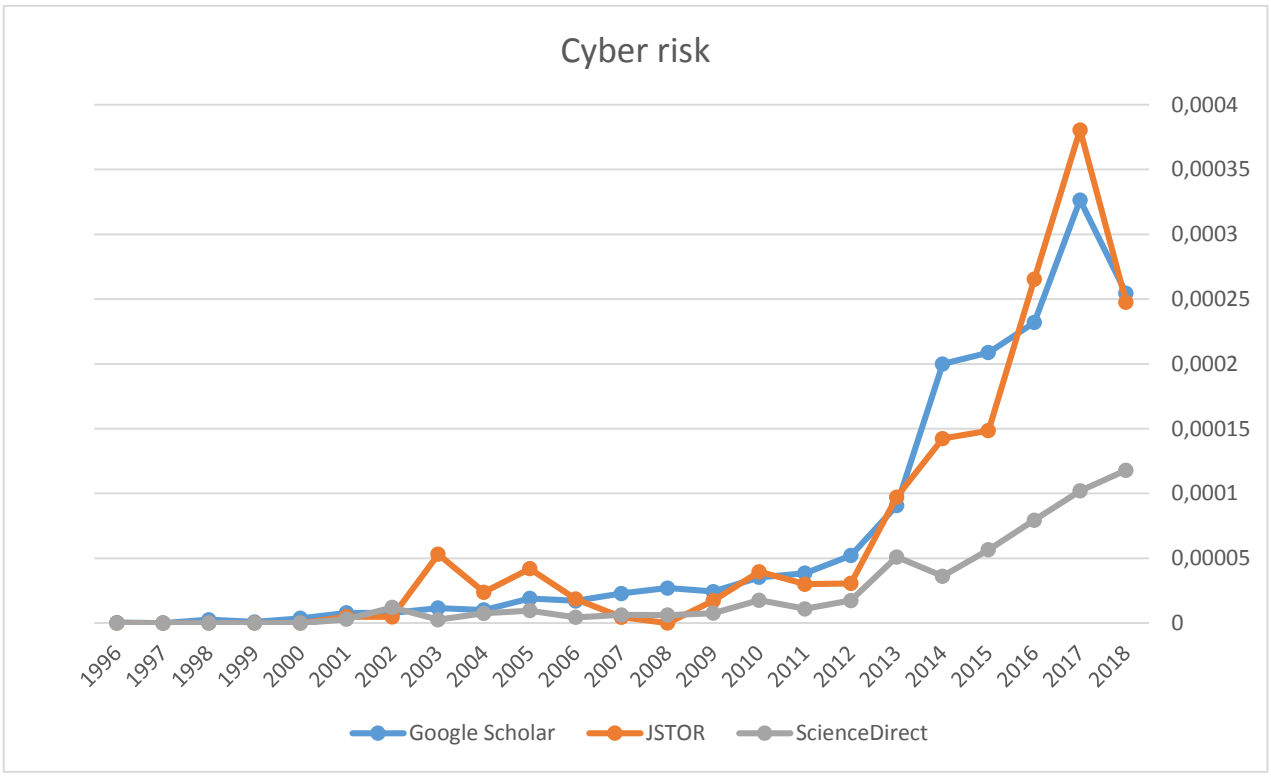

Рисунок 2.13 – Динаміка популярності біграми “Cyber risk”

На основі графіків наведених у рисунка 2.10-2.13 можемо говорити про різке зростання популярності терміну Кіберпростір починаючи з 1995 року його

певну фіксацію починаючи з 2000 року та новий ріст популярності у 2014 році. Терміни Кібербезпека, Кіберзагроза та Кібер ризик, починають активно вживатись в районі 2010-2012 років та досі знаходяться на етапі встановлення.

Тобто результати повністю співпадають з аналізом наведеним у пункті 2.2. І в подальшому при аналізі даних наведених у таблицях 2.1-2.5 можемо говорити саме про популярність термінології.

Відносно низьку популярність термінології за даними сервісу ScienceDirect можна пояснити його більшим спрямуванням на природничі, медичні науки та соціальні і гуманітарні науки.

2.4 Порівняльний аналіз

Для зображення розподілу вибірок побудовано графік “Ящик з вусами” для кожного з сервісів. Графіки побудовані за даними 2000-2018 років коли кожен з аналізованих термінів уже надійшов до вжитку.

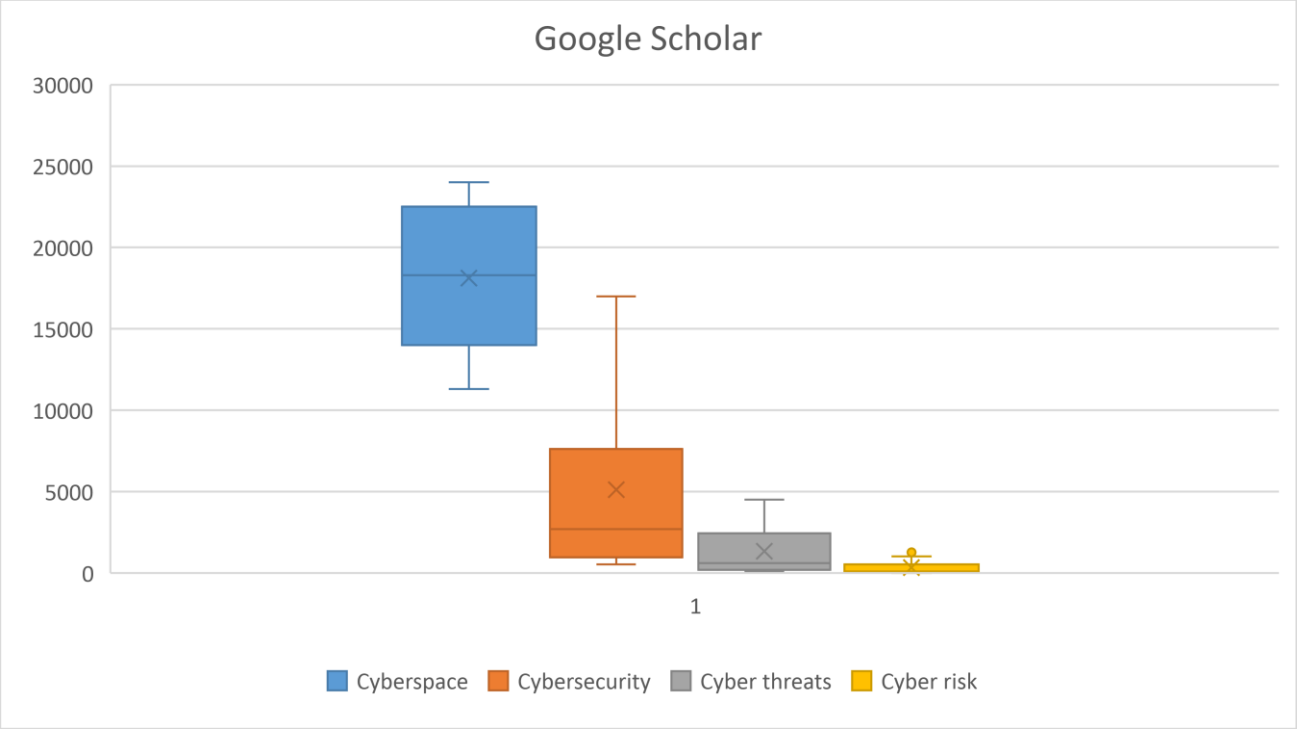

Рисунок 2.14 – “Ящик з вусами” за даними Google Scholar

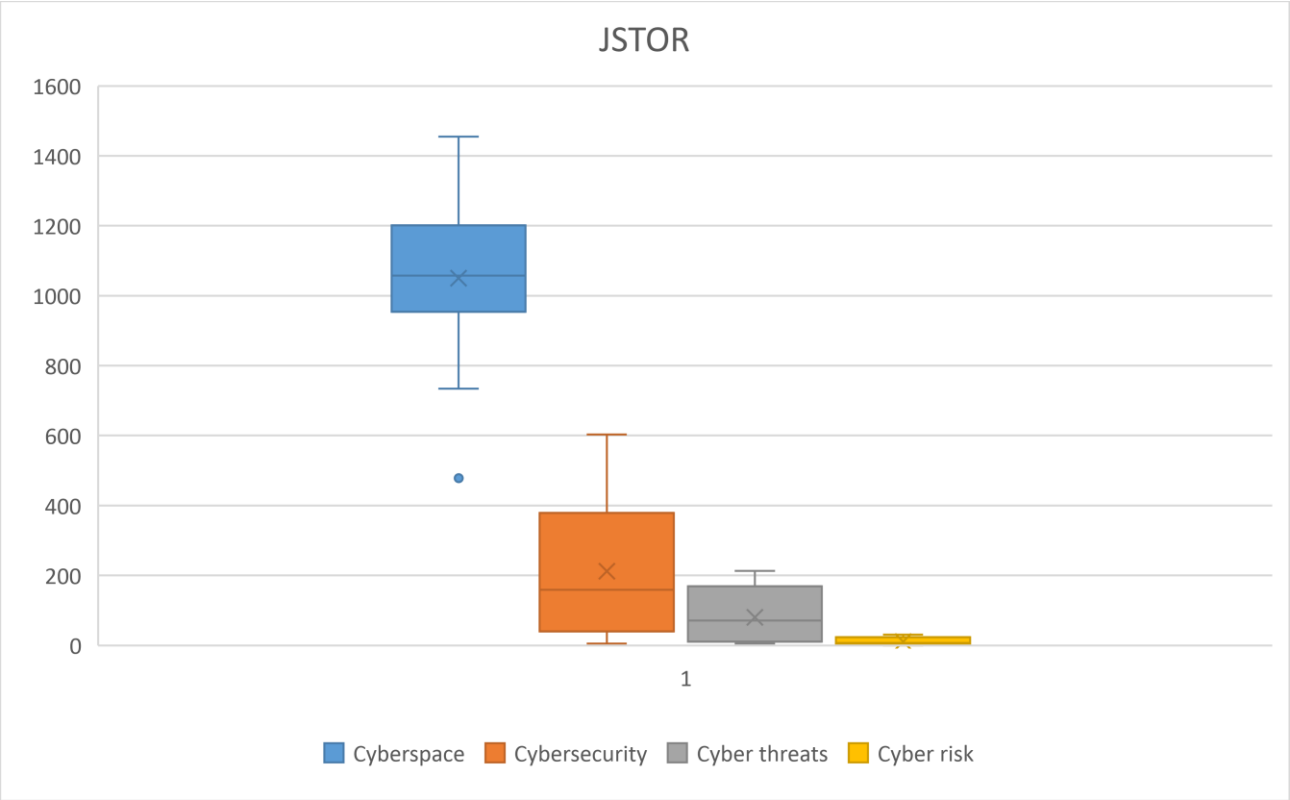

Рисунок 2.15 – “Ящик з вусами” за даними JSTOR

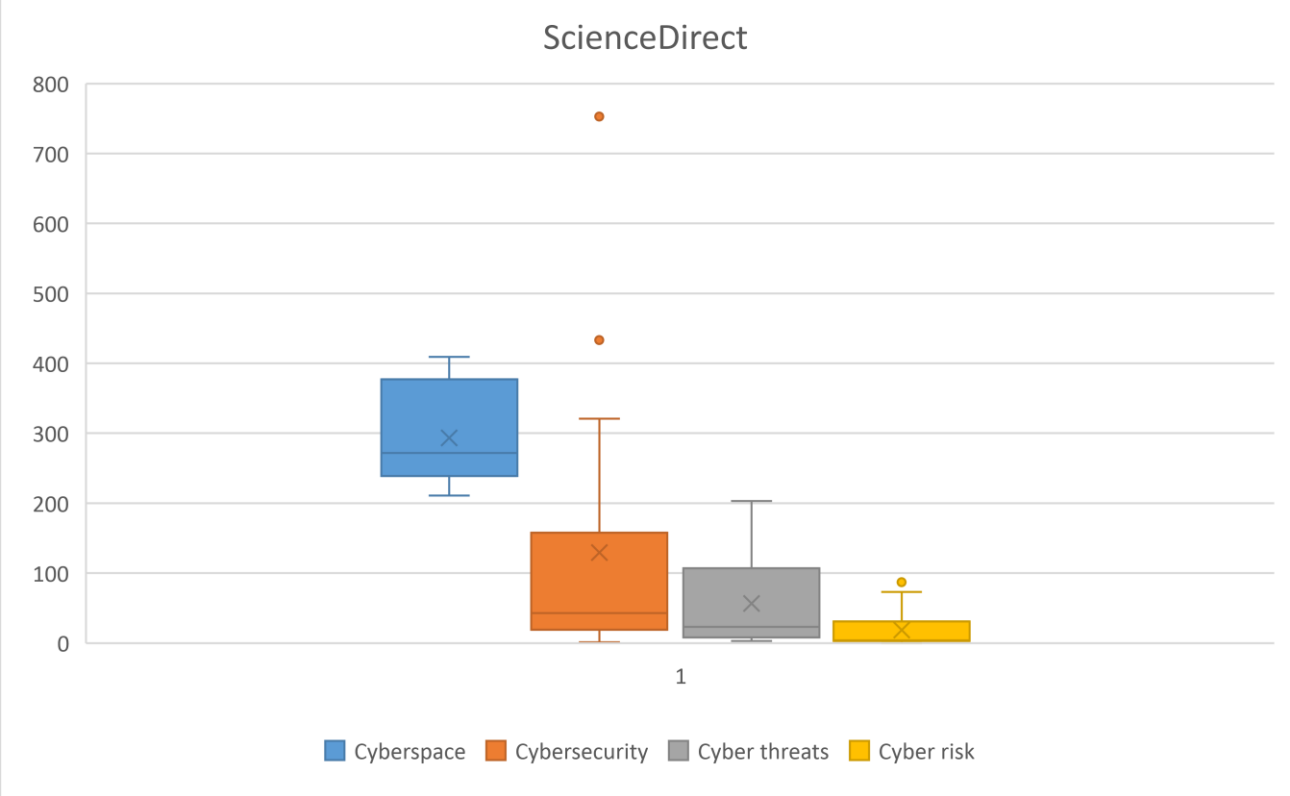

Рисунок 2.16 – “Ящик з вусами” за даними ScienceDirect

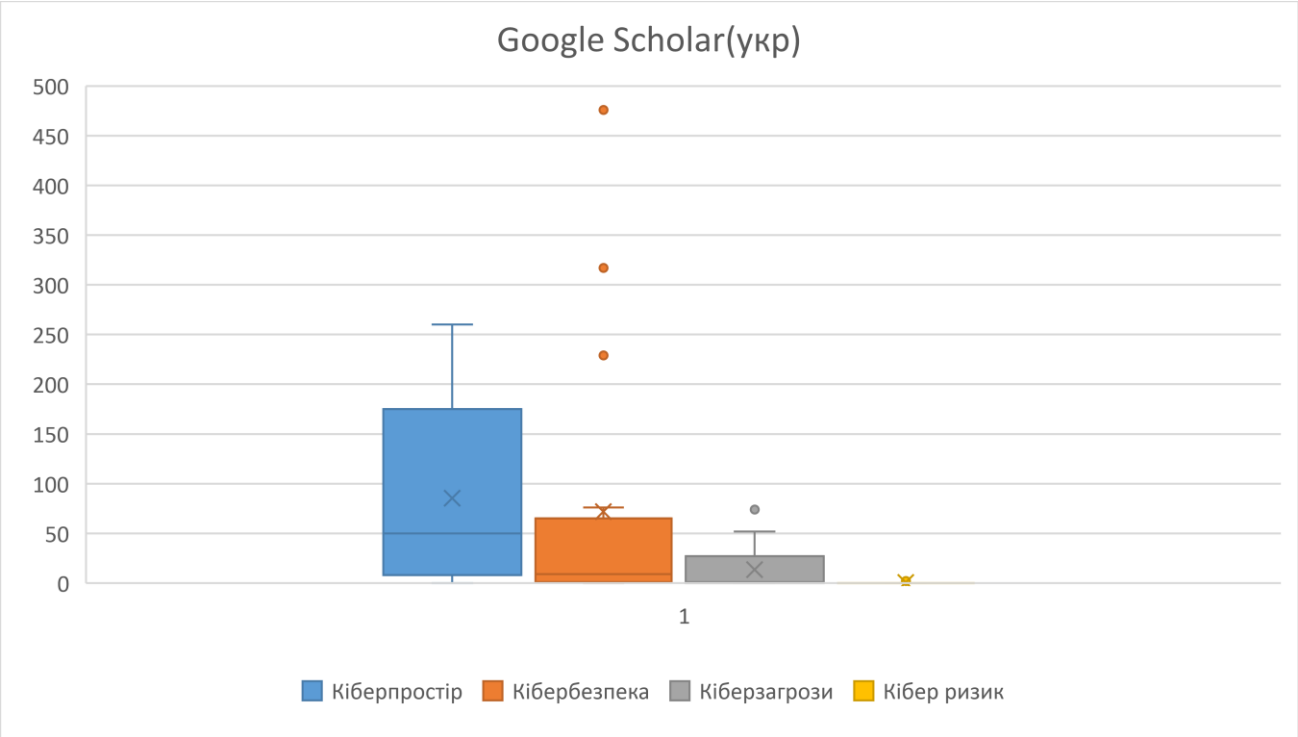

Рисунок 2.17 – “Ящик з вусами” за даними Google Scholar для україномовної літератури

Також для термінів “Cyberspace”, “Cybersecurity”, “Cyber threats”, “Cyber risk” усіх додатково було обраховано середнє арифметичне, медіану та середнє квадратичне за трьома наборами даних. Результати наведені у таблиці 2.8.

Таблиця 2.8 - Показники для порівняльного аналізу

Ресурс	Величина	Cyberspace	Cybersecurity	Cyber threats	Cyber risk
Google Scholar	Середнє	18115,79	5124,00	1346,68	331,05
	Медіана	18300,00	2710,00	623,00	108,00
	Середнє квадратичне відхилення	4396,12	5696,35	1465,12	428,44
Jstor	Середнє	1050,05	212,05	80,42	11,00
	Медіана	1057,00	159,00	71,00	7,00
	Середнє квадратичне відхилення	234,75	194,23	73,95	10,60

Продовження таблиці 2.8

Ресурс	Величина	Cyberspace	Cybersecurity	Cyber threats	Cyber risk
Sciencedirect	Середнє	293,37	129,58	56,74	18,79
	Медіана	272,00	43,00	23,00	4,00
	Середнє квадратичне відхилення	67,05	191,58	65,77	26,30

За середніми значеннями проаналізованих вибірок можна зробити висновок, що термін “Cyberspace” є найбільш вживаним. Причому порівняно з іншими розрив у декілька разів. Кібер ризики (Cyber risk) мають дуже низьку кількість цитувань у науковий літературі, оскільки використання цієї біграми можлива лише в вузько направленій літературі. Втім проблеми в цій сфері зараз знаходяться на зараз знаходяться на порядку денному тому Кіберзагрози(Cyber threats) та Кібербезпека(Cybersecurity) мають відносно велику кількість вживань.

Досить великі показники дисперсії свідчать про значні зміни в вживань термінів за розглянутий період. Дійсно, за розглянутий період спостерігалось інтенсивне зростання кількості наукових робіт в яких використовуються ці терміни, наразі найшивидше зростає вживання терміну “Cybersecurity”.

В україномовній літературі можемо спостерігати аналогічний результат, за виключенням того що термін “Кіберпростір” також знаходиться на стадії розвитку.

2.5 Аналіз використання n-грамм у друкованих джерелах в загалому

Для роботи було використано сервіс **Google Ngram Viewer** або **Google Books Ngram Viewer**[32] який дозволяє будувати графіки частотності N-грам на основі величезної кількості друкованих джерел, опублікованих з 16 століття і зібраних в сервіс Google Books. Для роботи з сервісом було проаналізовано публікації "Characterizing the Google Books Corpus: Strong Limits to Inferences of Socio-Cultural and Linguistic Evolution"[33], "The Pitfalls of Using Google Ngram to Study Language"[34] та опрацьовано книгу “Uncharted: Big Data as a Lens on Human Culture Riverhead Books”[35] від авторів цього сервісу Жана-Батист Мишеля та Ерец Ейдена.

Нажаль на даний момент простежити частотність можливо лише до 2008 року. Втім цього також достатньо щоб виявити тендінції. Як можемо бачити з Рисунку 2.18 дані отримані за допомогою Google Ngram Viewer підтверджують порівняльний аналіз наведений у попередньому пункті. А саме, серед загальної літератури найбільшого розповсюдження набув термін cyberspace (кіберпростір) який масово почав використовуватись у 1994 році та досяг піку популярності (за даними цієї вибірки) у 2000 році.

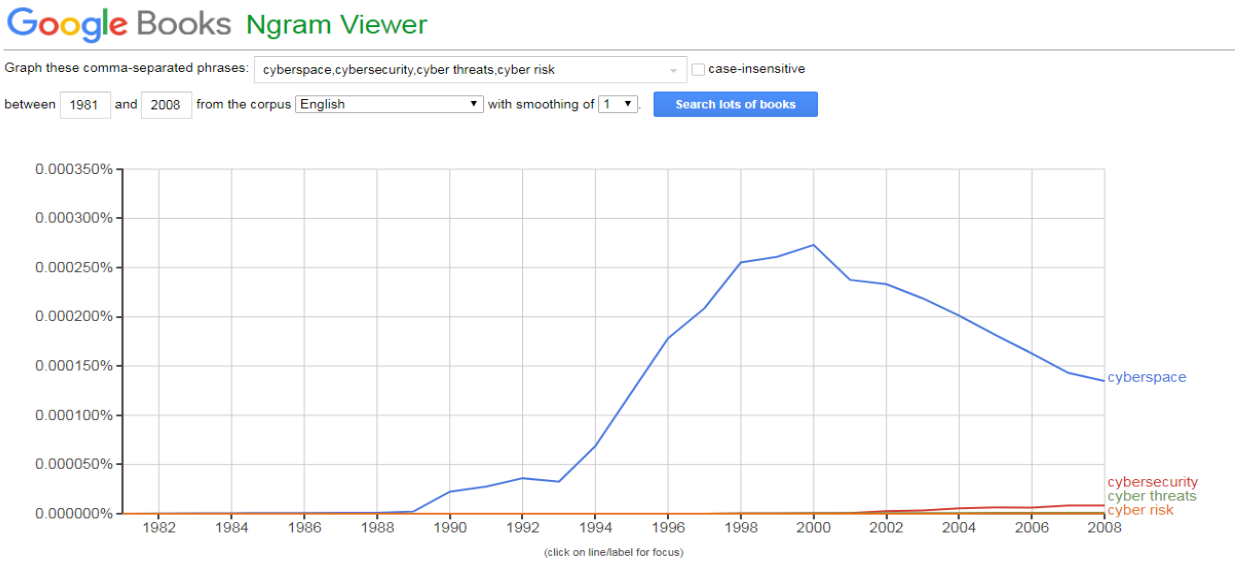

Рисунок 2.18 – Графік частости використання основних категорій серед друкованих джерел за 1981-2008 роки.

На рисунку 2.19 більш детально бачимо частоту використання варіантів терміну “Cyberspace” нечутливо до регістру.

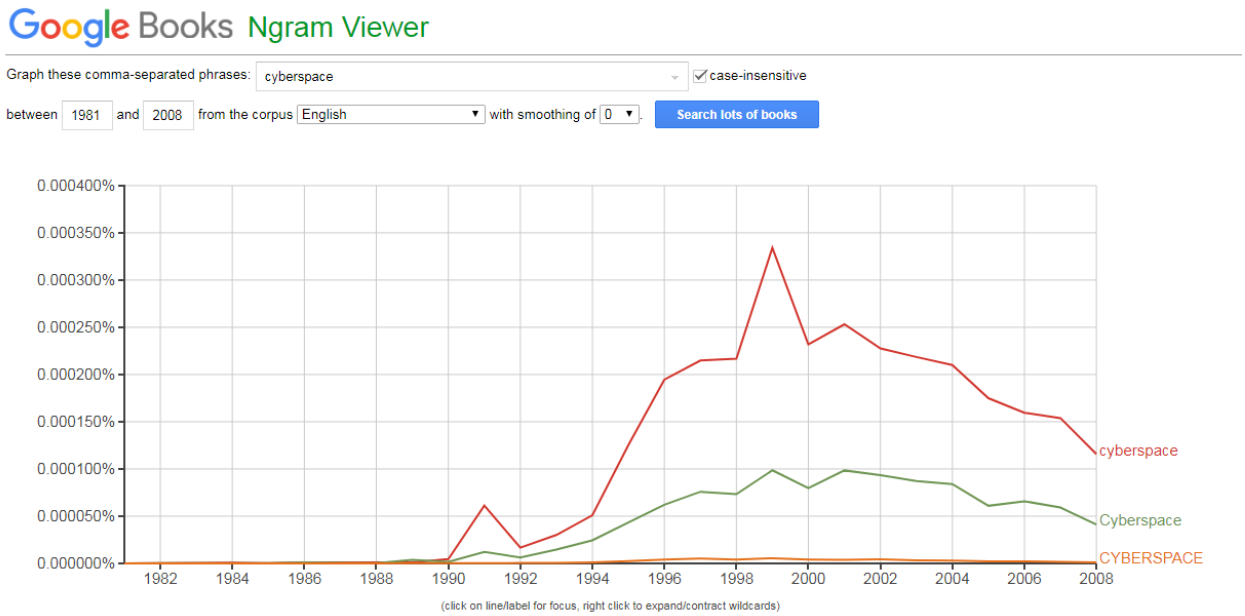

Рисунок 2.19 – Графік частоти вживання терміну “Кіберпростір” нечутливо до регістру.

З графіку наведеного у рисунку 2.20 бачимо що варіанти терміну “Cybersecurity” (Кібербезпека) почали активно застосовуватись з 2000 року та продовжують набирати популярність.

Рисунок 2.20 – Графік частоти вживання терміну “Кібербезпека” нечутливо до регістру.

В англомовній літературі досі відсутні однозначні правила правопису терміну “Кіберзагроза”. Тому на рисунку 2.21 наведено частотність використання можливих варіантів у множині та однині. Як можемо спостерігати термін почав активно вживатись з 1999 року і також стрімко набирає популярності. Найбільш часто використовуваною є біграма “cyber threats”

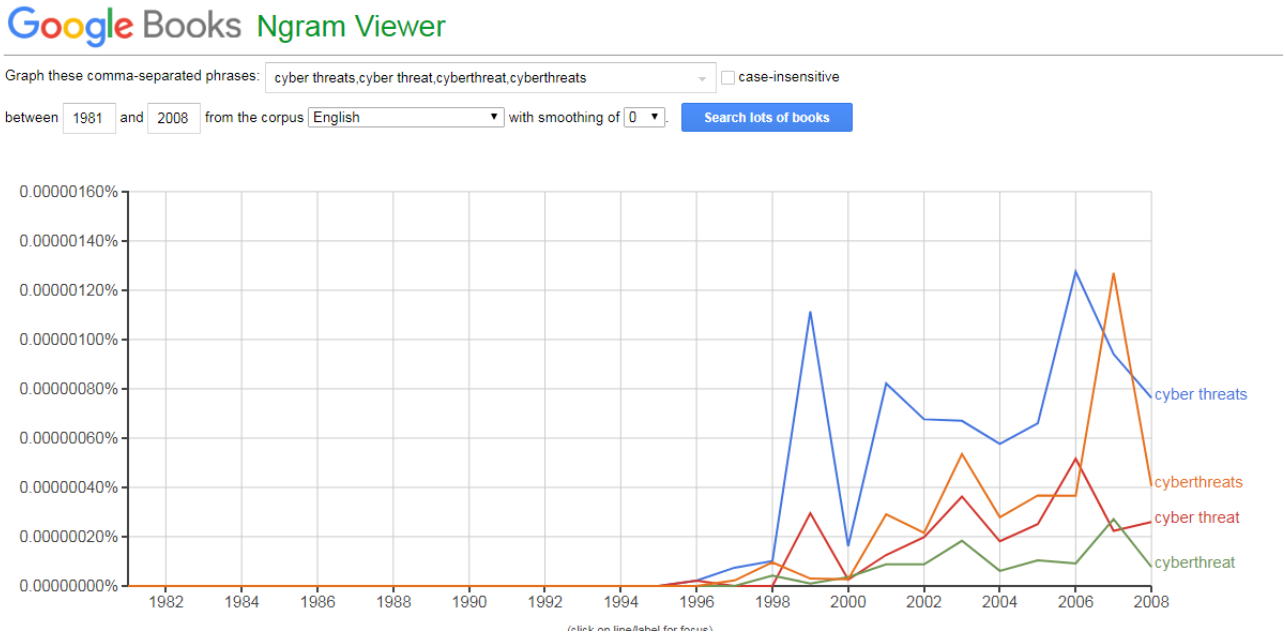

Рисунок 2.21 – Графік частотності вживання варіантів терміну “Кіберзагроза”

Термін “Кібер ризик” також досі знаходиться на етапі встановлення тому частотність можливих варіантів зображена на рисунку 2.22. Як бачимо варіанти “Cyberrisk” або “Cyberrisks” не є частими у вжитку. Найбільш часто термін використовуються у вигляді біграм, зокрема в однині - “Cyber risk”.Термін почав відносно часто з’являтись в літературі з 2003 року але досі має низьку популярність.

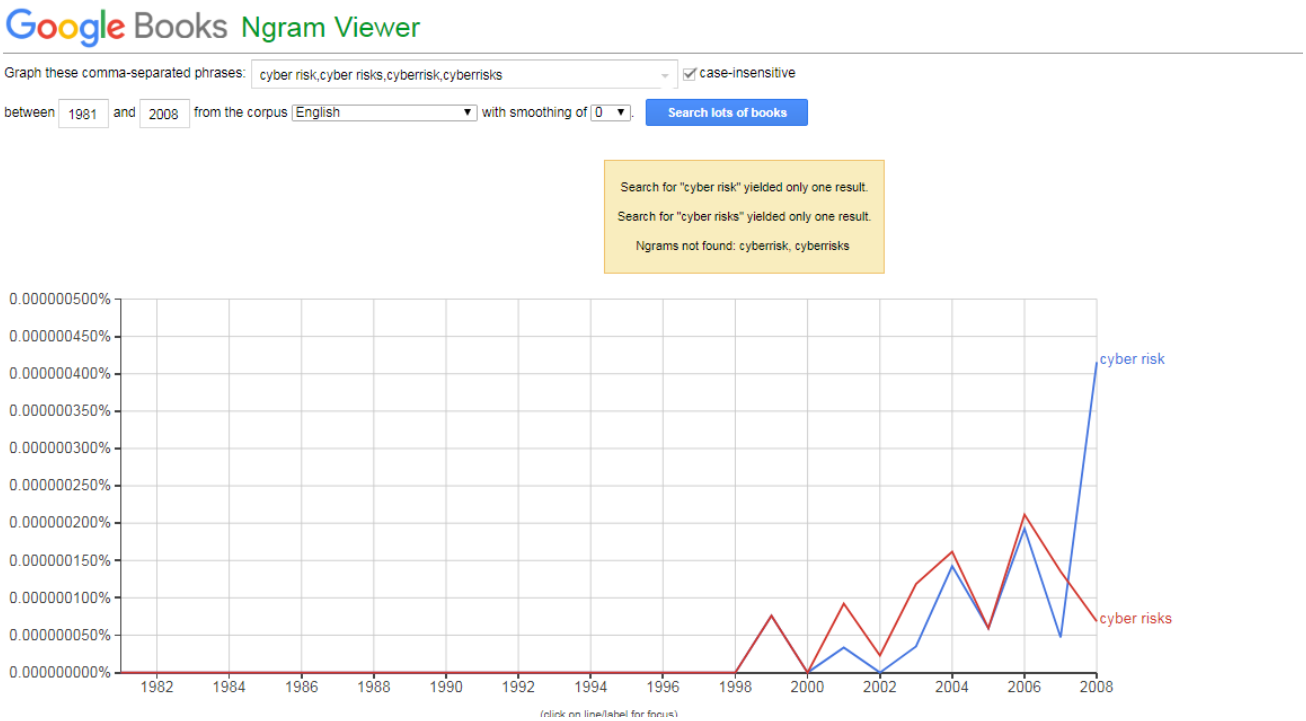

Рисунок 2.22 – Графік частотності вживання варіантів терміну “Кібер ризик”

Для усіх досліджуваних термінів можемо спостерігати значну похибку вимірювань у 1979 році(рисунок 2.23). Ресурс показує найчастіше використання термінів у цьому році проте жодних матеріалів з використанням шуканих термінів за даний рік знайти не вдалося.

Похибка виникає через включення великої кількості неправильно датованих і категоризованих текстів у базу даних Сервісу. Google Ngram Viewer також критикували через те що дані сервісу опираються на неточне оптичне розпізнавання символів та надлишок наукової літератури[36]. Втім для порівняння з результатами отриманими за науковим публікаціям вважаю використання цього інструменту доцільним.

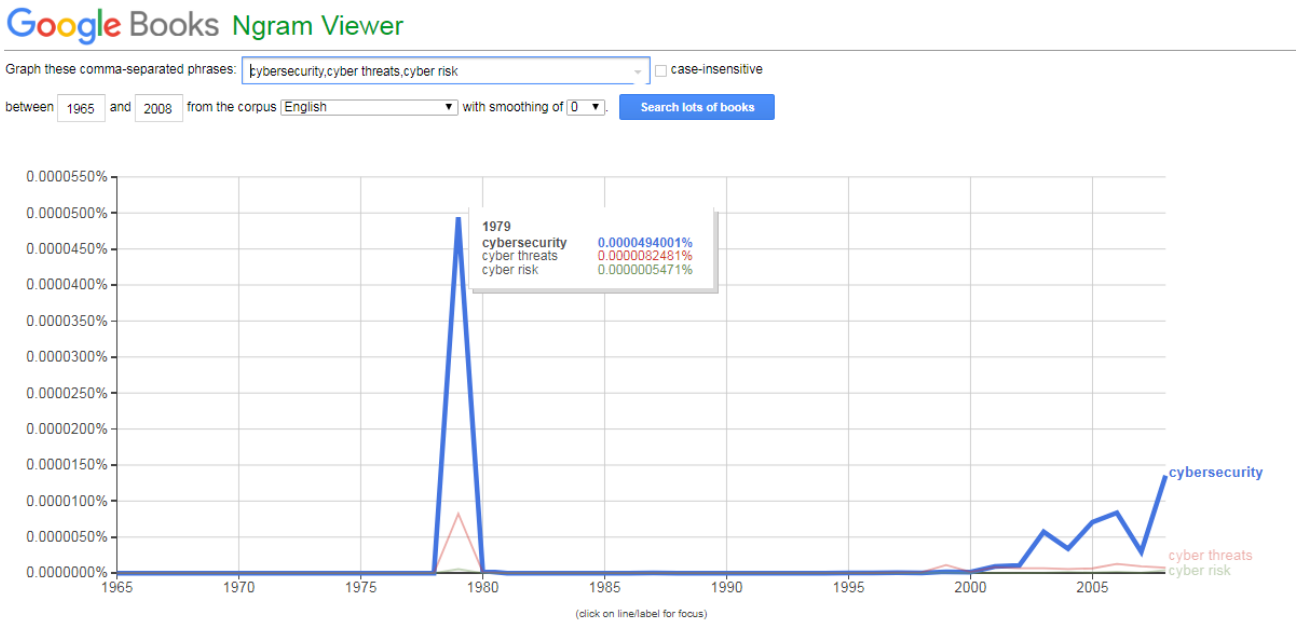

Рисунок 2.23 – Викид у даних сервісу Google Ngram Viewer

Висновки до розділу 2

У цьому розділі було проаналізовано динаміку вживання визначених елементів термінології у наукових роботах. Визначено що терміни надійшли до вжитку відносно недавно та їх популярність стрімко зростає. Перевірено отримані дані на кореляцію з динамікою росту наукових публікацій в цілому – кореляція присутня але повної залежності немає.

Також було проведено порівняльний аналіз між елементами термінології. Його результати та розвідкового аналізу підтверджуються даними з сервісу Google Ngram Viewer для друкованих джерел в цілому.

3 АНАЛІЗ ФРАКТАЛЬНОЇ СТРУКТУРИ ЧИСЛОВИХ РЯДІВ

Для аналізу наших систем використаємо запропонований Херстом [37] метод нормованого розмаху (RS-аналіз). Цей метод дозволяє розрізнити випадковий і фрактальний тимчасові ряди, а також робити висновки про наявність неперіодичних циклів, довготривалої пам'яті і т.д.

Розрахунок проводився за алгоритмом:

1) S_t – вихідний ряд. Розраховуємо логарифмічні відношення:
$$N_t = \ln\frac{S_t}{S_t - 1}$$

2) Розділяємо ряд N на A суміжних періодів довжини n. Кожен період позначаємо як I_a, де a = 1,2,…, A. Для кожного I_a визначаємо середнє значення:
$$E(I_a) = \frac{1}{n}\sum_{k=1}^{n} N_{k,a}$$

3) Для кожного періоду I_a розраховуємо відхилення від середнього значення:
$$X_{k,a} = \sum_{i=1}^{k}(N_{i,a} - E(I_a))$$

4) Розрахуємо розмах в межах кожного періоду:
$$R_{I_a} = \max\left(X_{k,a}\right) - \min\left(X_{k,a}\right)$$

5) Обраховуємо стандартне відхилення для кожного періоду I_a:
$$S_{I_a} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(N_{k,a} - E(I_a)\right)^2}$$

6) Кожен R_{I_a} ділимо на S_{I_a}. Потім розраховуємо середнє значення R/S:

$$R/S(n) = \frac{\sum_{a=1}^{A} R/S(A)}{A}$$

7) Збільшуємо n і повторюємо кроки 2-6 до тих пір, поки $n \leq N/2$

8) Будуємо графік залежності $\log\frac{R}{S(n)}$ від $\log n$ та за допомогою методу найменших квадратів знаходимо регресію виду: $\log\frac{R}{S(n)} = H \cdot \log n + c$, де H - показник Херста.

Обрахунки проводились за даними з сервісу Google Scholar оскільки він має найбільшу вибірку та допустимі показники кореляці (пункт 2.3). У таблиці 3.1 наведено результати роботи алгоритму для динаміки вживання терміну “Кіберпростір”.

Таблиця 3.1 – Результати R/S аналізу вживань терміну “Cyberspace”

n	$R/S(n)$	$\log n$	$\log R/S(n)$
3	1,386945183	0,477121255	0,142059296
4	1,643347428	0,602059991	0,215729389
5	2,003102322	0,698970004	0,301703134
6	1,972708645	0,77815125	0,295062948
7	2,409433241	0,84509804	0,381914898
8	2,75022874	0,903089987	0,439368816
9	3,26042682	0,954242509	0,513274457
10	3,748434425	1	0,573849918
11	3,85556872	1,041392685	0,586088448
12	4,226650186	1,079181246	0,625996305
13	5,600936702	1,113943352	0,748260665
14	5,74004296	1,146128036	0,758915143
15	5,995613977	1,176091259	0,777833663
16	5,95518346	1,204119983	0,774895145
17	5,804922946	1,230448921	0,763796459
18	5,48748228	1,255272505	0,739373131

На основі результатів наведених в таблиці 3.1 побудовано графік залежності $\log\frac{R}{S(n)}$ від $\log n$ зображений на рисунку 3.1.

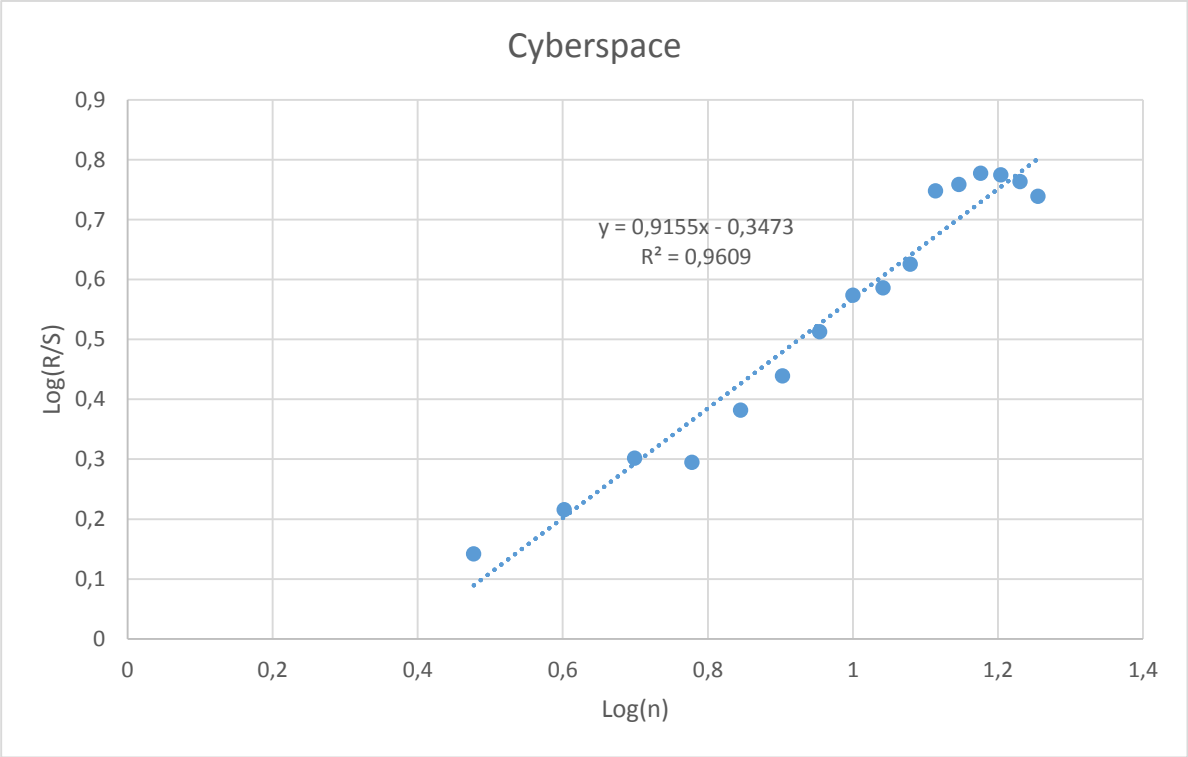

Рисунок 3.1 – Графік залежності $\log\frac{R}{S(n)}$ від $\log n$ для динаміки вживань терміну “Кіберпростір”

За допомогою методу найменших квадратів побудовано регресію, яка дорівнює y = 0,9155x - 0,3473 де 0,9155 це показник Херста(H). Як було доведено у роботі[38] для випадкового процесу з незалежними приростами і скінченною дисперсією H дорівнює 0,5. Також у роботах[39-40] на великому емпіричному матеріалі було встановлено що для природних процесів показник Херста групується навколо значень 0,72-0,73.

В даній же роботі на основі праці[41] показник Херста використовується як індекс залежності("index of long-range dependence"). Тобто відмінність показника Херста від 0.5 є відображенням фрактальних властивостей процесів.

Значення H в діапазоні 0,5–1 вказує часовий ряд з довгостроковою позитивною автокореляцією, а при значенні H < 0,5 тенденція змінюється на

протилежну (зростання спостерігається величини змінюється спаданням і навпаки).

Значення $H = 0{,}5$ може вказувати на абсолютно некорельований ряд, але на практиці це значення часто вказує на ряди для яких автокореляція на малих числових вибірках можуть бути позитивними або негативними, але абсолютні значення автокореляцій швидко спадають до нуля.

Для динаміки вживань терміну “Кіберпростір” маємо H=0,9155, тобто послідовність персистентна і процес буде зберігати наявну тенденцію.

Далі наведено результати для трьох інших категорій:

Таблиця 3.2 – Результати R/S аналізу вживань терміну “Cybersecurity”

n	$R/S(n)$	$\log n$	$\log R/S(n)$
3	1,34468514	0,47712125	0,128620606
4	1,607404033	0,60205999	0,206125054
5	1,863862703	0,69897	0,270413918
6	2,194320305	0,77815125	0,341300022
7	2,533689245	0,84509804	0,403753348
8	3,021565805	0,90308999	0,480232057
9	2,998053236	0,95424251	0,47683934
10	2,998053236	1	0,47683934
11	2,69462649	1,04139269	0,430498575

Таблиця 3.3 – Результати R/S аналізу вживань біграми “Cyber threats”

n	$R/S(n)$	$\log n$	$\log R/S(n)$
3	1,3726	0,4771	0,137534546
4	1,6934	0,6021	0,228761952
5	2,1062	0,699	0,323501777
6	1,8524	0,7782	0,267735281
7	2,3779	0,8451	0,376186006
8	2,7847	0,9031	0,444772744
9	3,3805	0,9542	0,528975397
10	3,3805	1	0,528975397
11	3,7772	1,0414	0,577164984

Таблиця 3.4 – Результати R/S аналізу вживань біграми “Cyber risk”

n	$R/S(n)$	$\log n$	$\log R/S(n)$
3	1,358673851	0,47712125	0,133115217
4	1,591693127	0,60205999	0,201859341
5	1,796699643	0,69897	0,254475482
6	1,902847245	0,77815125	0,279403926
7	2,120246649	0,84509804	0,326386386
8	2,56783987	0,90308999	0,409567938
9	2,3069755	0,95424251	0,363042982
10	2,3069755	1	0,363042982
11	2,850568686	1,04139269	0,45493151

На основі результатів наведених в таблицях 3.2-3.4 за допомогою методу найменших квадратів побудовано лініні регресії наведені на рисунках 3.2-3.4

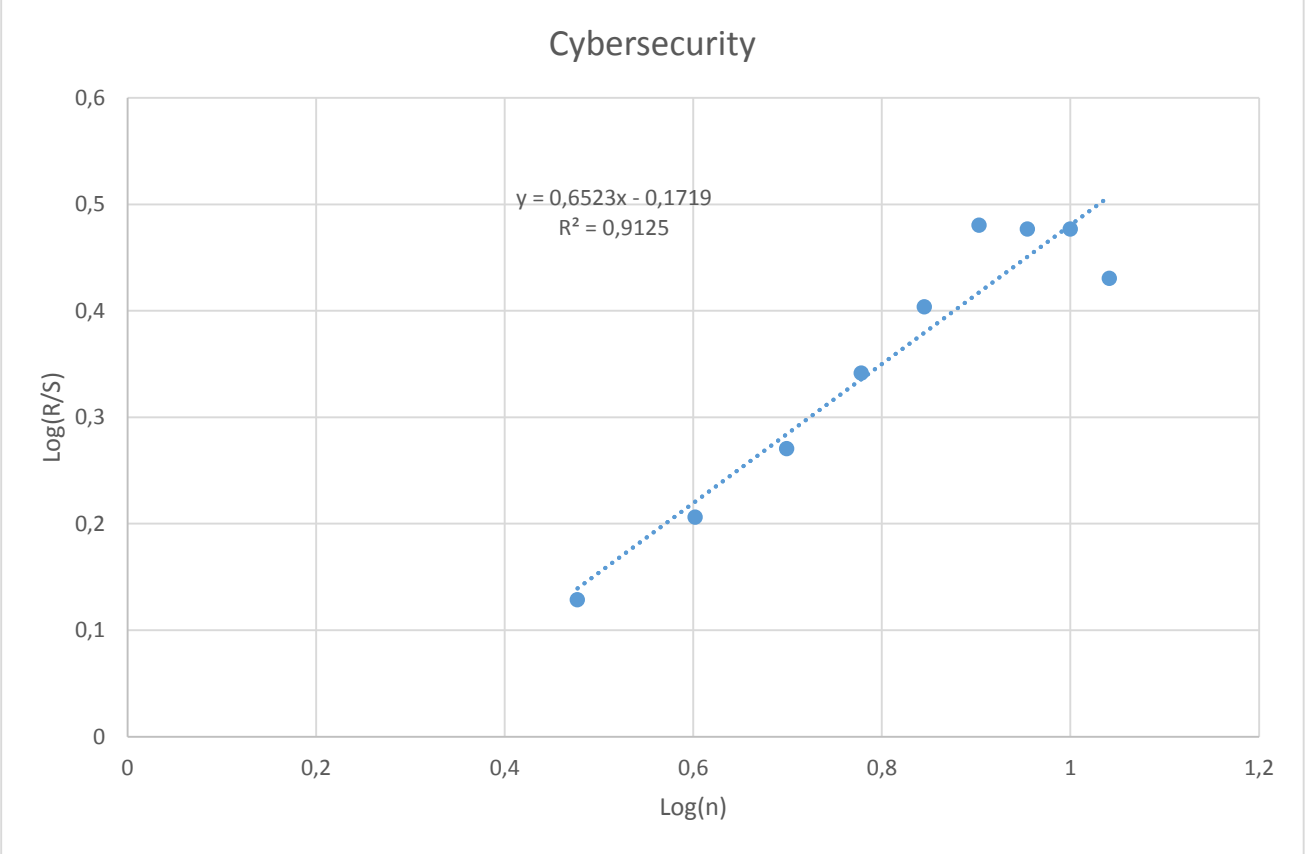

Рисунок 3.2 – Графік залежності $\log\frac{R}{S(n)}$ від $\log n$ для динаміки вживань терміну “Кібербезпека”

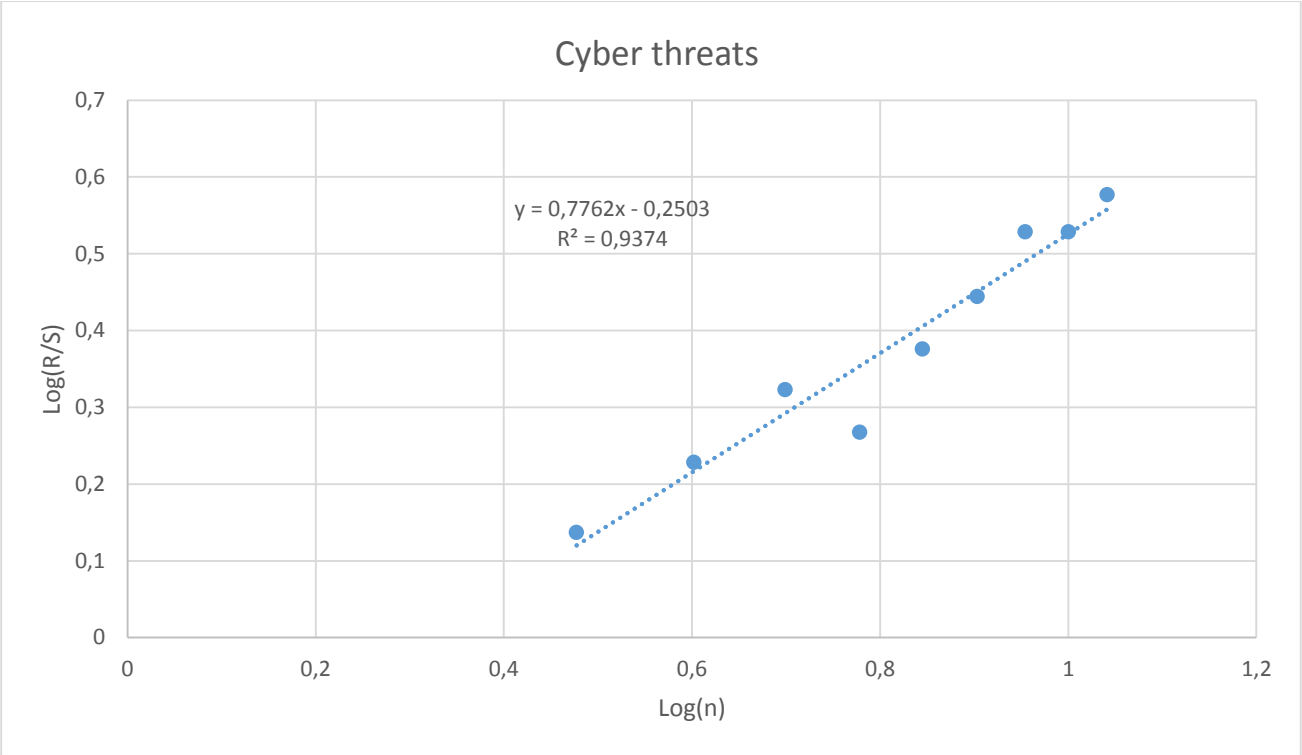

Рисунок 3.3 – Графік залежності $\log \frac{R}{S(n)}$ від $\log n$ для динаміки вживань терміну “Кіберзагроза”

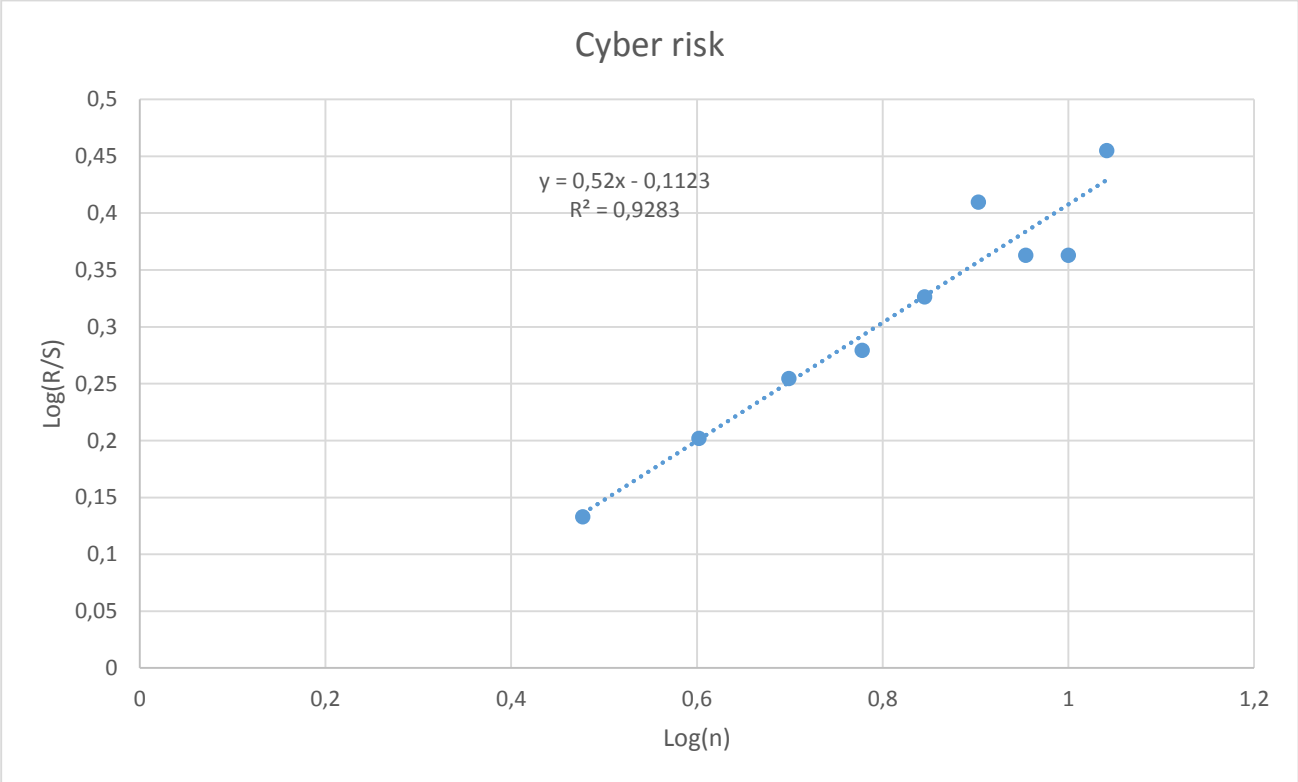

Рисунок 3.4 – Графік залежності $\log \frac{R}{S(n)}$ від $\log n$ для динаміки вживань терміну “Кібер ризик”

Відповідно маємо коефіцієнти Хешера для динаміки вживань “Cybersecurity” H = 0,6523, для “Cyber threats” H=0,7762 та “Cyber risk” H= 0,52.

Висновки до розділу 3

Тобто маємо ряди з довгострокою позитивною автокореляцію для вживань термінів “Кіберпростір”, “Кібербезпека”, “Кіберзагроза” що говорить про те що тенденції зростання популярності цієї термінології виявлені у другому пункті будуть зберігатись.

За результатами R/S аналізу динаміка вживання біграми “Кібер ризик” носить випадковий характер і відповідає звичайному гаусівському шуму. Це пов’язано з тим що термін почав активно вживатись тільки у осанні декілька років тому даний алгоритм не може простежити тенденцію.

4 ПОРІВНЯННЯ ВИКОРИСТАННЯ БЕЗПЕКОВОЇ ТЕРМІНОЛОГІЇ В ІНФОРМАЦІЙНІЙ ТА КІБЕРНЕТИЧНИХ СФЕРАХ

Терміни «кібербезпека» і «інформаційна безпека» зазвичай використовуються як синоніми в термінології безпеки і створюють плутанину серед фахівців з безпеки.

Термінологічні відмінності між цими сферами проаналізовано в роботах [1-2] та буде звернено увагу у роботі Сьомака Р.В. “Системний аналіз категорії “загроза” в інформаційній безпеці”[42].

В цьому плані додатково можна зазначити ключову відмінність у цих сферах а саме, те що: кібербезпека займається захистом систем управління а інформація безпека – захистом інформації.

На основі даних наведених у таблицях 2.1-2.4 та даних по вживанню категорій інформаційної безпеки які будуть опубліковані в роботі[42] для проведення порівняльного аналізу побудовано графіки “Ящик з вусами” по кожній з категорій.

Графіки побудовані за даними 1996-2018 років коли термінологія у сфері кібербезпеки надійшла до вжитку. У даній роботі на рисунках 4.1-4.4 зображено графіки на основі даних сервісу Google Scholar оскільки він має найбільшу вибірку. Результати додатково отримані за допомогою сервісів ScienceDirect та JSTOR співпадають.

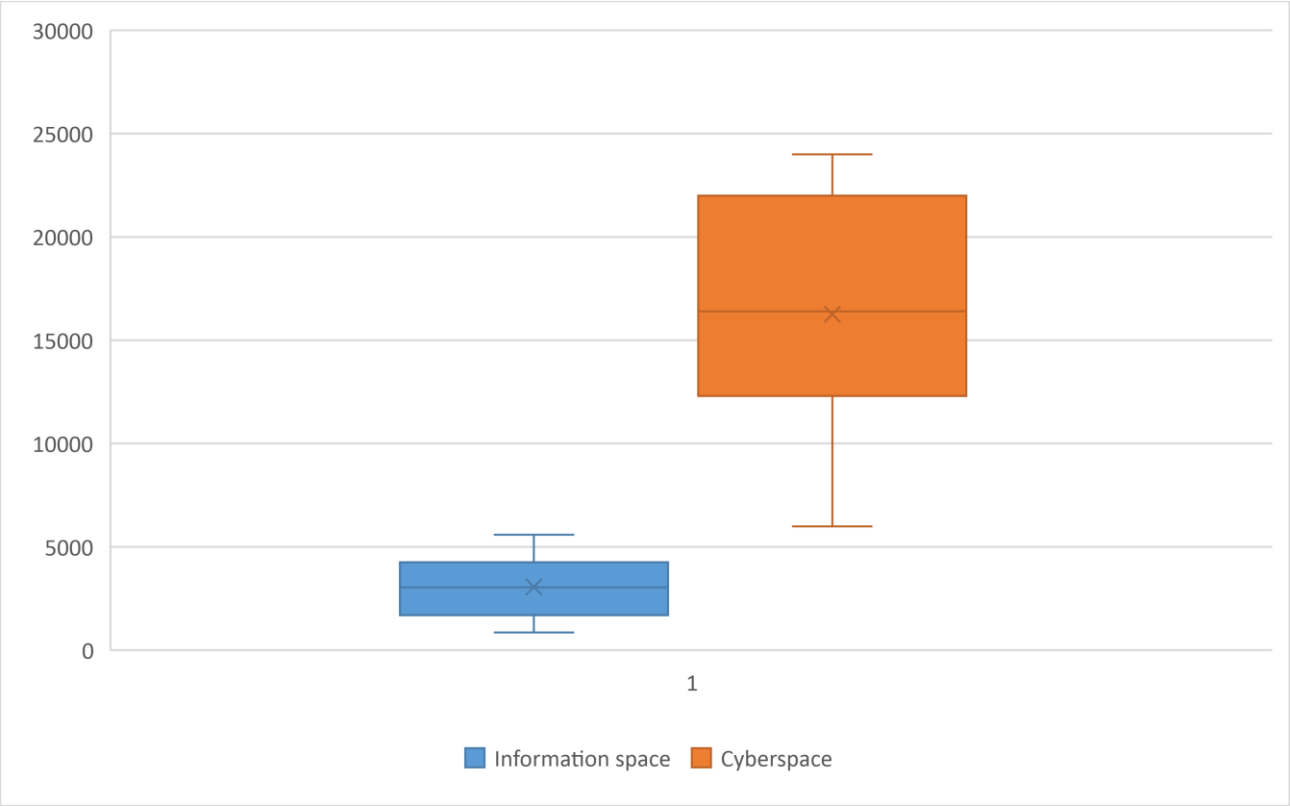

Рисунок 4.1 – “Ящик з вусами” для термінів “Простір”

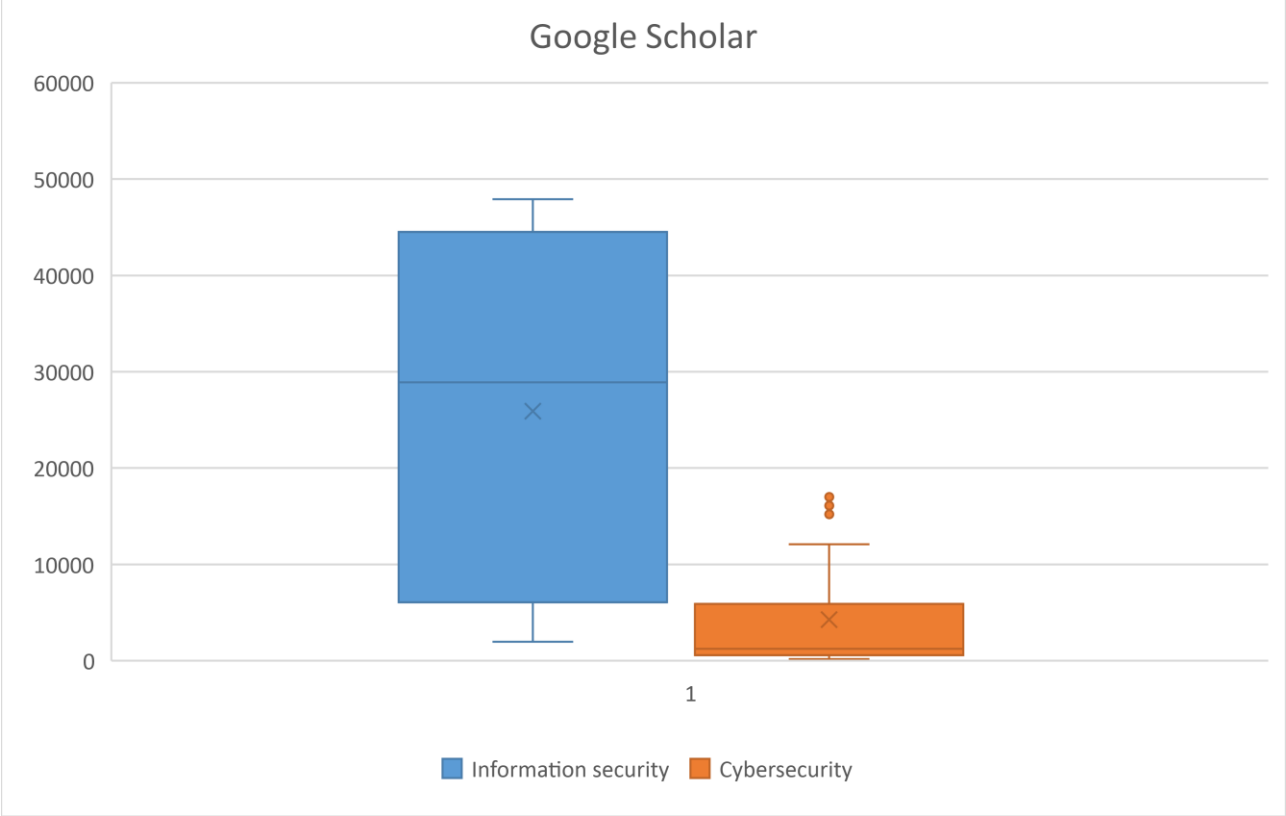

Рисунок 4.2 – “Ящик з вусами” для термінів “Безпека”

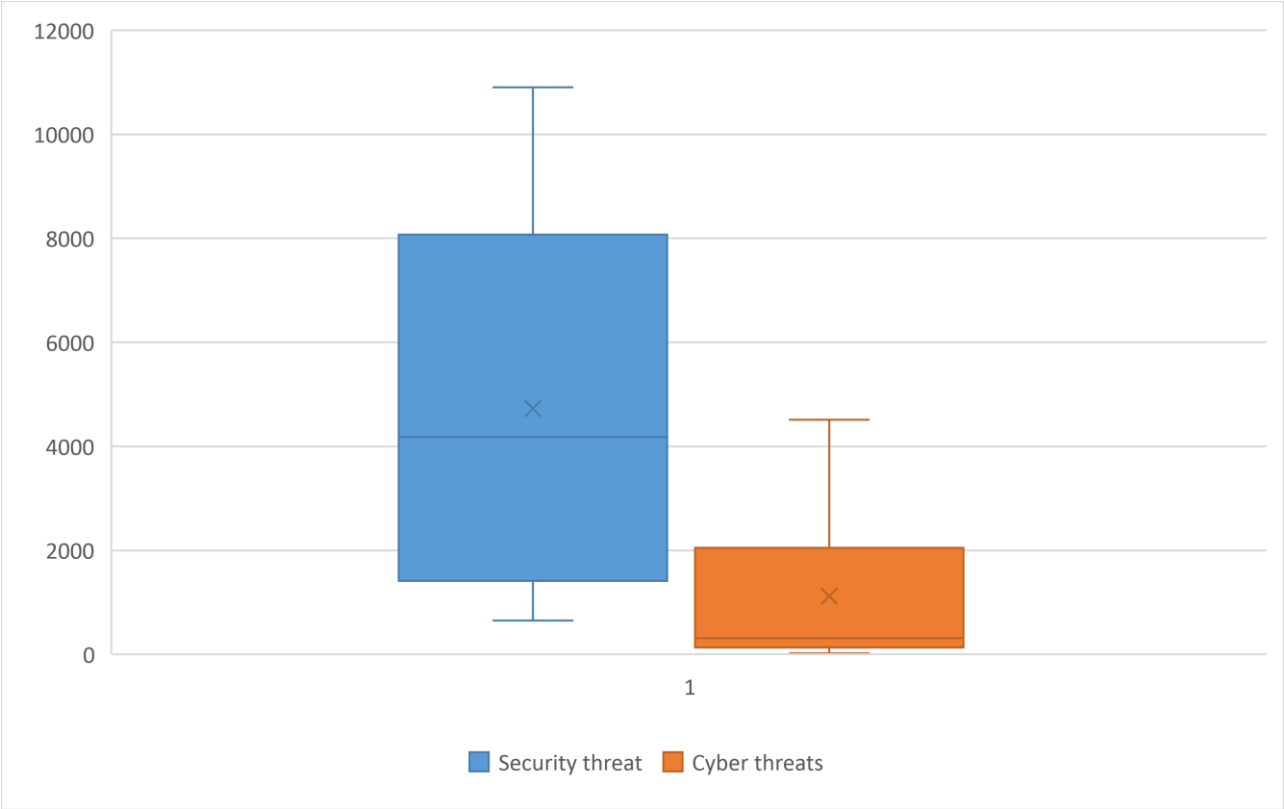

Рисунок 4.3 – “Ящик з вусами” для термінів “Загроза”

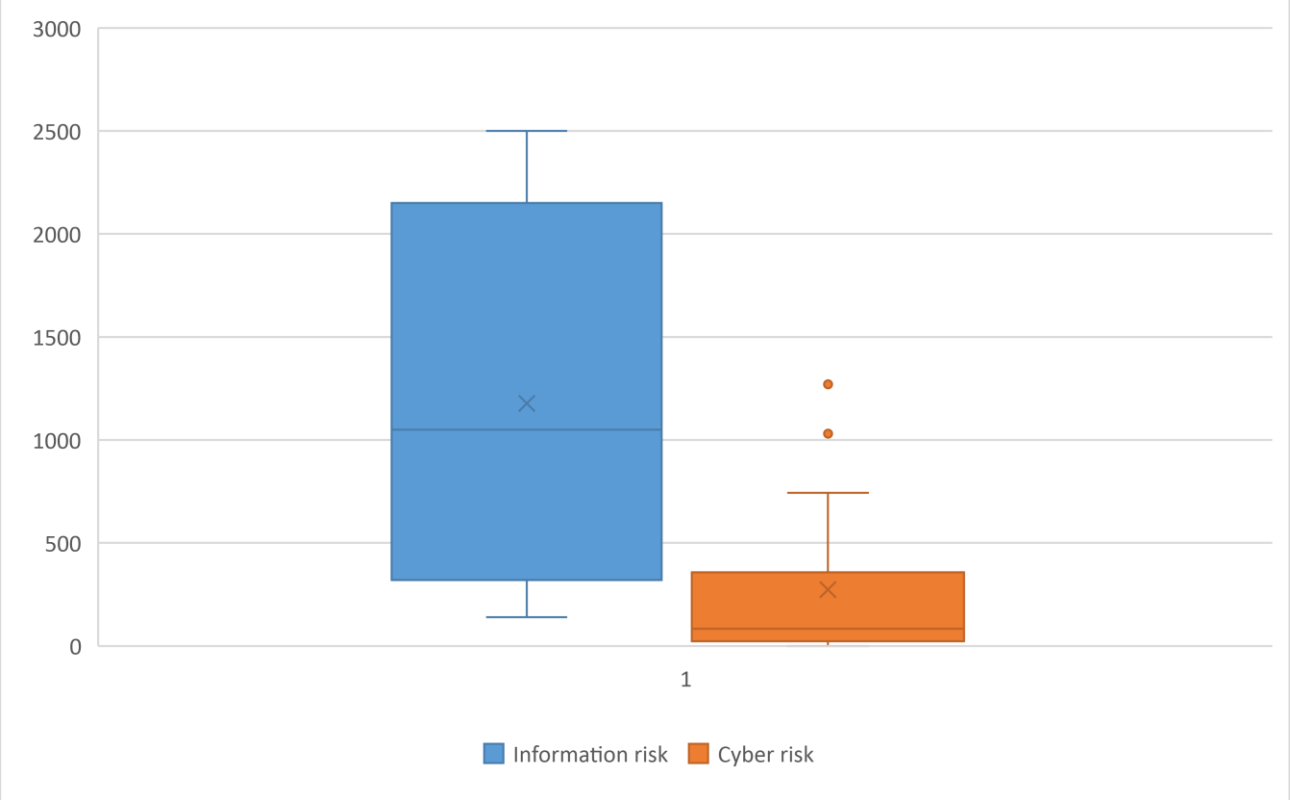

Рисунок 4.4 – “Ящик з вусами” для термінів “Ризик”

Як можемо бачити термін “Кіберпростір” у наукових публікаціях популярніший за “Іформаційний простір” у 5 разів. Дискусія у безпековій сфері(безпека, ризик, загроза) навпаки вдвічі більше ведеться про інформаційну безпеку порівняно з кібербезпекою.

У таблиці 4.1 обраховано кореляцяю між динамікою вживань у наукових роботах термілогії в інформаційній та кібер сфері.

Кореляцію розраховано за проміжком з 1996 по 2018 роки за формулою:

$$r_{xy} = \frac{\sum(x_i - x_{\text{сер}})(y_i - y_{\text{сер}})}{\sqrt{\sum(x_i - x_{\text{сер}})^2 * \sum(y_i - y_{\text{сер}})^2}}$$

Таблиця 4.1 – Коефіцієнти кореляції кількості наукових робіт у конкретні роки з вживанням розглянутих елементів кібер та інформаційної сфери.

Сервіс	Простір	Безпека	Загроза	Ризик
Google Scholar	0,954	0,653	0,901	0,790
ScienceDirect	0,338	0,544	0,951	0,768
JSTOR	0,583	0,592	0,732	-0,128

Дані залежно від сервісу значно відрізняються, це пов’язано з різними спрямуваннями робіт у них. Зокрема ScienceDirect більше спрямований на природні та гуманітарні науки, тому термінологія в сфері інформаційної безпеки рідко зустрічається у ньому. Втім за останні 2 роки там з’явилась значна кількість робіт пов’язаних з кібербезпекою.

Тому найбільш релевантним виглядає Google Scholar через включення у себе робіт всіх напрямів та найбільший обсяг вибірки. За результати кореляції по даним цього сервісу можемо стверджувати що зростання популярності термінів “Кіберпростір”, “Кіберзагроза”, “Кібер ризик” має майже повну позитивну залежність з зростанням популярності термінів “Інформаційний простір”, “Інформаційна загроза” та “Інформаційний ризик”.

Вживання ж термінів “Кібербезпека” та “Інформаціна безпека” немає такої чітко вираженої кореляції. І в останні 5 років “Кібербезпека” значно швидше набирає популярність порівняно з “Інформаціною безпекою”.

Висновки до розділу 4

У даному розділі проведено порівняльний аналіз динаміки обсягу наукової дискусії в сферах інформаційної безпеки та кібербезпеки. Також обраховано кореляцію між двома цими процесами.

ВИСНОВКИ

У ході цієї роботи було детально проаналізовано термінологічну проблему в контексті загрози у кібербезпеці. Встановлено ключові елементи, а саме:

1. Кіберпростір (Cyberspace)
2. Кібербезпека (Cybersecurity)
3. Кіберзагрози (Cyber threats)
4. Кібернетичний ризик (Cyber risk)

Розглянуто визначення цих термінів які використовуються у основних нормативно-правових документах західних країн та в роботах присвячених кібербезпеці. На основі проведеного аналізу виокремлено основні елементи необідні для однозначного трактування цих термінів та запропоновані власні визначення.

Проаналізовано динаміку використання зазначеної термінології у наукових роботах та у друкованих джерелах в цілому. Оскільки розглянуті процеси не є нормально-розподіленими застосовано R/S аналіз для виявлення фрактальної структури, неперіодичних циклів та довготривалої пам'яті. Результати аналізу дозволили зробити прогноз подальшої динаміки вживання термінології.

Проведено порівняльний аналіз динаміки обсягу наукової дискусії в сферах інформаційної безпеки та кібербезпеки.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Качинський А.Б. Безпека складних систем.-К. [Текст]: ТОВ “Видавництво “Юстон”, 2017.-498 с

2. Грайворонський М.В. Сучасні підходи до забезпечення кібернетичної безпеки [Текст] / Матеріали XVII Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених “Теоретичні і прикладні проблеми фізики, математики та інформатики”, НТУУ “КПІ”, 2015 р.

3. Про основні засади забезпечення кібербезпеки України, Закон України. [Електронний ресурс] – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/2163-19

4. Oxford Dictionaries [Електронний ресурс] – Режим доступу до ресурсу: https://en.oxforddictionaries.com/definition/us/cyberspace

5. Gibson W. Neuromancer [Текст] / W. Gibson.— London: HarperCollins, 1994

6. The varieties of cyberspace: Problems in definition and delimitation, Western Journal of Communication, 63:3, 382-412. [Електронний ресурс] – Режим доступу до ресурсу: https://www.tandfonline.com/doi/abs/10.1080/10570319909374648

7. "Cyberpower and National Security: Policy Recommendations for a Strategic Framework," in Cyberpower and National Security, FD Kramer, S. Starr, [Текст] L.K. Wentz (ed.), National Defense University Press, Washington (DC) 2009.

8. THE ITU NATIONAL CYBERSECURITY STRATEGY GUIDE [Текст] - Dr. Frederick Wamala (Ph.D.), CISSP®- September 2011

9. ISO/IEC 27032:2012 [Текст] Information technology — Security techniques — Guidelines for cybersecurity

10. Department of Defense “Dictionary of Military and Associated Terms”, [Текст] 12 April 2001 (As Amended Through 17 October 2007)

11. Department of Defense “Dictionary of Military and Associated Terms”, [Текст] 12 April 2001 (As Amended Through 19 August 2009)

12. Department of Defense “Dictionary of Military and Associated Terms”, [Текст] 8 November 2010 (As Amended Through 15 June 2013)

13. DECREE OF THE PRESIDENT OF THE COUNCIL OF MINISTERS (DPCM) [Текст] January 24, 2013 Directive laying down guidelines for cybernetic protection and the national cyber security.

14. Ministry of Public Safety, [Текст] “Canada’s Cyber Security Strategy,” 2018 ст.33-34

15. Federal Ministry of the Interior, [Текст] “Cyber Security Strategy for Germany Report,” 2018 ст. 14-15

16. MED, “NEW ZEALAND’S Cyber Security Strategy,” [Текст] p. 3, 2015

17. HM Government, “NATIONAL CYBER SECURITY STRATEGY 2016-2021” [Текст] p.75 2016.

18. Joint Staff, “DOD Dictionary of Military and Associated Terms,” [Текст] 2018. ст. 60

19. National Cybersecurity Awareness Month, [Текст] 2009

20. Computers & Security, Volume 15, Issue 3, 1996, Page 221 . Автор Helen Meyer. [Електронний ресурс] – Режим доступу до ресурсу: https://www.sciencedirect.com/journal/computers-and-security

21. Oxford Dictionaries [Електронний ресурс] – Режим доступу до ресурсу: https://en.oxforddictionaries.com/definition/us/cybersecurity

22. “Inside the Army Vol. 8, No. 24 (June 17, 1996), pp. 1, 11-12 ” автор Jason Sherman. [Електронний ресурс] – Режим доступу до ресурсу: https://www.jstor.org/stable/43979220?seq=1#page_scan_tab_contents

23. Oxford Dictionaries [Електронний ресурс] – Режим доступу до ресурсу: https://en.oxforddictionaries.com/definition/cyberthreat

24. “Risks of Cyber Attack to Water Utility Supervisory Control and Data Acquisition Systems” Military Operations Research, Vol. 6, No. 2, Special Issue: Information Operations (2001), pp. 23-33 автори Barry C. Ezell, Yacov Y. Haimes and James H. Lambert. [Електронний ресурс] – Режим доступу до ресурсу: https://www.jstor.org/stable/43943673?seq=1#page_scan_tab_contents

25. Michigan Law Review Vol. 93, No. 8 (Aug., 1995), pp. 2440-2447 [Електронний ресурс] – Режим доступу до ресурсу: https://www.jstor.org/stable/1289940?seq=1#page_scan_tab_contents

26. Google Scholar [Електронний ресурс] – Режим доступу до ресурсу: https://scholar.google.com

27. Springer Link [Електронний ресурс] – Режим доступу до ресурсу: https://link.springer.com/article/10.1007%2Fs11192-018-2958-5

28. Plosone [Електронний ресурс] – Режим доступу до ресурсу: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0093949

29. JSTOR [Електронний ресурс] – Режим доступу до ресурсу: https://www.jstor.org/

30. ScienceDirect [Електронний ресурс] – Режим доступу до ресурсу: https://www.sciencedirect.com/

31. Nature [Електронний ресурс] – Режим доступу до ресурсу: https://www.nature.com/articles/d41586-018-07841-9

32. Google Ngram Viewer [Електронний ресурс] – Режим доступу до ресурсу: https://books.google.com/ngrams

33. Pechenick, Eitan Adam; Danforth, Christopher M.; Dodds, Peter Sheridan; Barrat, Alain (7 October 2015). "Characterizing the Google Books Corpus: Strong Limits to Inferences of Socio-Cultural and Linguistic Evolution". PLOS ONE.

10 (10): e0137041. [Електронний ресурс] – Режим доступу до ресурсу: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0137041

34. Zhang, Sarah. "The Pitfalls of Using Google Ngram to Study Language". WIRED. Retrieved 2017-05-24. [Електронний ресурс] – Режим доступу до ресурсу: https://www.wired.com/2015/10/pitfalls-of-studying-language-with-google-ngram/

35. Uncharted: Big Data as a Lens on Human Culture Riverhead Books; [Текст] 1st edition (December 26, 2013)

36. Nunberg, Geoff (16 December 2010). "Humanities research with the Google Books corpus". [Електронний ресурс] – Режим доступу до ресурсу: https://web.archive.org/web/20160310035741/http://languagelog.ldc.upenn.edu/nll/?p=2847

37. Херст, Г. Э., 1951. «Долгосрочная вместимость водохранилищ». [Текст] Труды Американского общества гражданских инженеров, 116, 770-808.

38. Feller W. The asymptotic distribution ofthe range ofsums ofindependent variables [Текст] // Ann. Math. Statist. 1951. V. 22. P. 427–432.

39. Федер Е. Фракталы. [Текст] М.: Мир, 1991.

40. Андреев С. Д., Ивлев Л. С. Временная и пространственная изменчивость полей оптических и аэрозольных характеристик в атмосфере. Ч. I. Оптические характеристики атмосферы [Текст] // Оптика атмосферы и океана. 1997. Т. 10, № 12. С. 1440–1449.

41. Torsten Kleinow (2002)Testing Continuous Time Models in Financial Markets, [Текст] Doctoral thesis, Berlin

42. “Системний аналіз категорії “загроза” в інформаційній безпеці” [Текст] Сьомак Р.В. 2019.